

Release 3.1A John F. Collins, Biocomputing Research Unit  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

```

Mpsrch_pp    protein - protein database search, using Smith-Waterman algorithm
Run on:      Thu Jun 22 21:23:22 2000;      MasPar time 4.06 Seconds
Tabular output not generated.               145.748 Million cell updates/sec

```

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect score: 183  
Sequence: 1 EEFLLTPKKLQCVDLHVISNDVCAQV 25

Scoring table: PAM 150  
Gap 15

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 08  
Listing first 45 summaries

Database: a-geneseq35  
1:geneseqF

Statistics: Mean 20.967; Variance 66.122; scale 0.317

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

No	Score	Query	Match	Length	DB	ID	Description	Pred. No.
1	183	100.0	237	1	W96187		Human prostate specific	3.20e-13
2	183	100.0	237	1	M83213		Prostate-specific ant1	3.20e-13
3	183	100.0	237	1	R77098		Prostate-specific ant1	3.20e-13
4	183	100.0	237	1	W56086		Human prostate specific	3.20e-13
5	183	100.0	250	1	W03130		Prostate-specific ant1	3.20e-13
6	183	100.0	261	1	M13649		Human prostatic specif	3.20e-13
7	178	97.3	237	1	R84671		Mature kallikrein hk3	1.34e-12
8	178	97.3	237	1	W45398		Prostate-specific ant1	9.87e-12
9	171	93.4	30	1	W27387		Prostate specific ant1	3.20e-07
10	134	73.2	240	1	W11023		Human prostate specif1	9.62e-07
11	130	71.0	237	1	W96186		Mature human Kallikrei	9.62e-07
12	130	71.0	237	1	W96190		Variant human Kallikrei	9.62e-07
13	130	71.0	237	1	R84667		Mature Kallikrein hk2	9.62e-07
14	130	71.0	237	1	W49087		Mutant human Kallikrei	9.62e-07
15	130	71.0	237	1	M83212		hk2 variant A217V	9.62e-07
16	130	71.0	237	1	W45395		Mature prostate-specif	9.62e-07
17	130	71.0	237	1	M83202		Prostate-specific glan	9.62e-07
18	130	71.0	238	1	R84670		Mature Kallikrein hk2	9.62e-07
19	130	71.0	244	1	W96188		Pro human Kallikrein 2	9.62e-07
20	130	71.0	244	1	M83204		Prostate-specific glan	9.62e-07
21	130	71.0	244	1	W45396		Prostate-specific glan	9.62e-07
22	130	71.0	244	1	R84669		Pro-hk2 kallikrein	9.62e-07
23	130	71.0	261	1	W06911		Prostate-specific glan	9.62e-07

25	130	71.0	261	1	R6468	Prepro-rh2 kallikrein.	9.62e-07
24	130	71.0	261	1	R6489	Prepro human kallikrel	9.62e-07
26	130	71.0	261	1	W4905	Wild-type human kallik	9.62e-07
27	130	71.0	261	1	W06972	Kallikrein prepro-hk2v	9.62e-07
28	130	71.0	261	1	W45397	Prostate-specific glan	9.62e-07
29	130	71.0	261	1	W10500	Human prepro-Typ226-gl	9.62e-07
30	130	71.0	261	1	W45400	Prostate-specific specifi	9.62e-07
31	130	71.0	261	1	W83203	Prostate-specific glan	9.62e-07
32	116	63.4	15	1	W58047	Human prostate specifi	4.36e-05
33	111	65.7	15	1	W58048	Human prostate specifi	1.67e-04
34	110	60.1	15	1	W58050	Human prostate specifi	2.19e-04
35	108	59.0	15	1	W58049	Human prostate specifi	3.73e-04
36	98	53.6	18	1	R55193	Residues 140-157 of th	5.23e-03
37	96	55.5	20	1	W11022	Antigen peptide derive	8.81e-03
38	93	50.8	15	1	W58046	Human prostate specifi	1.932e-02
39	91	49.7	245	1	P70677	Human kallikrein gene	3.22e-02
40	91	49.7	262	1	P70568	Human kallikrein-like	3.22e-02
41	91	49.7	262	1	P55121	Kallikrein encoded by	3.22e-02
42	91	49.7	262	1	W71005	Human prostate-associa	3.22e-02
43	90	49.2	15	1	W58051	Human prostate specifi	4.16e-02
44	90	49.2	17	1	W50784	Prostate Specific Antl	4.16e-02
45	88	48.1	245	1	P22314	Human recombinant kall	6.94e-02

## ALIGNMENTS

[illegible]

RESULT 2  
ID W83213 standard; Protein; 237 AA.  
AC W83213:  
DT 09-FEB-1999 (first entry)  
DE Prostate-specific antigen protein hK3.  
KW pphk2; prostate-specific antigen; hK2; hK1; hK3; pphk2; tissue kallikrein;  
KW pphk2; prostate-specific glandular kallikrein protein; PSA; human;  
KW Prostate-specific antigen.  
OS Homo sapiens.  
PN W09846795-A1.  
PD 22-OCT-1998.  
PF 09-APR-1998; U07027.  
PR 11-APR-1997; US-843076.  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
PI (MAYO-) MAYO FOUNDATION.  
PI Slawin KM, Tindall DJ, Young CYF;  
DR WPI: 98-594592/50.  
PT Detection of human kallikrein 2 RNA - by reverse transcription and  
PT amplification by PCR, for detecting, monitoring and staging of  
PT prostate cancer  
PS Disclosure: Fig 1: 90pp: English.  
CC The present invention describes a diagnostic method for detecting human  
CC kallikrein 2 (hk2) DNA. The method comprises: (a) contacting DNA obtained  
CC by reverse transcription (RT) of RNA from a human physiological sample  
CC which comprises cells suspected of containing hk2 RNA with at least 2  
CC oligonucleotides to amplify the DNA by PCR to yield amplified hk2 DNA,  
CC where the conditions amplify the DNA obtained by RT of RNA from at least  
CC one cell containing hk2 in a sample which comprises at least 107 to 109  
CC cells; and (b) detecting the presence of the amplified hk2 DNA. The  
CC method can be used for detecting, monitoring the progression of and  
CC pathologically staging prostate cancer. The present sequence represents  
CC prostate-specific antigen protein hK3.  
SQ Sequence 237 AA:

Query Match 100.0%; Score 183; DB 1; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3.20e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 139 EEFLLPKKLCQVDLHVSNVCAOV 163  
QY 1 EEFLLPKKLCQVDLHVSNVCAOV 25

RESULT 3  
ID R77098 standard; Protein; 237 AA.  
AC R77098:  
DT 16-JAN-1996 (first entry)  
DE Prostate-specific antigen.  
KW Prostate-specific antigen; prostate cancer; diagnosis.  
OS Homo sapiens.  
PN W09528498-A1.  
PD 26-OCT-1995.  
PF 14-APR-1995; U04680.  
PR 15-APR-1994; US-229391.  
PA (UYCO ) UNIV COLUMBIA NEW YORK.  
PI Buttyan R, Katz AE, Olsson CA, Ratfo A;  
DR WPI: 95-373812/48.  
DR N-PSDB: T04864.  
PT Method for enhancing prostate-specific antigen detection - provides  
PT sensitive means to identify early stages of prostate cancer.  
PS Disclosure: Page 69-70; 94pp: English.  
CC A prostate-specific antigen (PSA) is detected by isolating mRNA from  
CC a sample, producing cDNA, and subjecting the cDNA to RT-PCR using  
CC primers specific for the human PSA coding region (given in T04864).  
SQ Sequence 237 AA:

Query Match 100.0%; Score 183; DB 1; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3.20e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 139 EEFLLPKKLCQVDLHVSNVCAOV 163  
QY 1 EEFLLPKKLCQVDLHVSNVCAOV 25

RESULT 4  
ID W56086 standard; Protein; 237 AA.  
AC W56086:  
DT 11-AUG-1998 (first entry)  
DE Human prostate specific antigen protein.  
KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
KW detection; cancer; serine protease.  
OS Homo sapiens.  
PN W09810292-A1.  
PD 12-MAR-1998.  
PF 25-AUG-1997; U14909.  
PR 06-SEP-1996; US-025404.  
PA (GENZ ) CENTOCOR INC.  
PI Heaven GA;  
DR WPI: 98-193789/17.  
PT Monoclonal antibodies specific for prostate specific antigen -  
PT useful, e.g. in screening for prostate or breast cancer and  
PT especially to distinguish between benign prostatic hyperplasia and  
PT prostate cancer  
PS Disclosure: Page 32-33; 84pp: English.  
CC The present sequence represents prostate specific antigen (PSA).  
CC Monoclonal antibodies specific for PSA and hydrolases producing them  
CC have been developed. The antibodies: (a) bind to free PSA; (b) are  
CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL)  
CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and  
CC 15.2, 156 or 225 binding to amino acids 139-144 (EEFLFP) and 55-60  
CC (SLRHPF) respectively of free and bound PSA, or fragments. The antibodies  
CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
CC comprising a solid support with attached labelled monoclonal antibody  
CC specific for free PSA (especially (b)) and a PSA standard can be used.  
CC To detect both free and bound PSA, a second solid support with attached  
CC (differently labelled) monoclonal antibody binding free and bound PSA  
CC (especially selected from (c)) can be used either with or in place of,  
CC the first solid support. The antibodies are useful in cancer screening,  
CC especially prostate and breast cancer. By obtaining total and free PSA  
CC values, their ratio can be used to separate prostatic cancer (Pca) from  
CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
CC useful after radical prostatectomy, to predict disease persistence. The  
CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
CC of BPH and Pca patients with PSA values 4-10 ng/ml not previously  
CC possible by total PSA testing, avoiding biopsies.  
SQ Sequence 237 AA:

Query Match 100.0%; Score 183; DB 1; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3.20e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 139 EEFLLPKKLCQVDLHVSNVCAOV 163  
QY 1 EEFLLPKKLCQVDLHVSNVCAOV 25

RESULT 5  
ID W03130 standard; Protein; 250 AA.  
AC W03130:  
DT 21-OCT-1996 (first entry)  
DE Prostate-specific antigen.  
KW Prostate-specific antigen; metastasis; carcinoma; tumour; cancer;  
KW neoplasm; archival tissue; polymerase chain reaction; RT-PCR;  
KW kallikrein.  
OS Homo sapiens.  
PN key  
FT key  
FT peptide  
FT 1.13  
FT /label= Pro-peptide  
FT 14..250  
FT /label= Mat-protein  
FT /note= "prostate-specific antigen"  
PN W09621042-A2.  
PD 11-JUL-1995  
PF 04-JAN-1996; U00461.  
PR 04-JAN-1995; US-368706.

PA (UYBO-) UNIV BOSTON.  
 PT Edelstein RA, Moreland RB;  
 DR WPI: 96-334017/33.  
 DR N-PSDB: 135867.  
 PT Detecting a metastatic disorder in a fixed biological sample  
 PT comprises reverse transcription PCR of sample nucleic acid and  
 PT detection of metastatic, esp. prostate, tissue-specific target  
 PT sequence  
 PS Disclosure: Fig 1; 60pp; English.  
 CC Prostate-specific antigen (PSA) (W03130) is a serine protease  
 CC (kallikrein) found in both normal and neoplastic prostate  
 CC epithelium. It can be used as a prostate marker. PCR primers  
 CC (see also 135855-62) based on PSA cDNA (135867) or genomic  
 CC (135868) sequences can be used for the specific amplification of PSA  
 CC nucleic acids in fresh and archival tissue samples as a means of  
 CC detecting metastatic sequences.  
 SQ Sequence 250 AA;

Query Match 100.0%; Score 183; DB 1; Length 230;  
 Best Local Similarity 100.0%; Pred. No. 3,20e-13;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 152 EEFLLPKKLCQVLDHVISNDVCAOV 176  
 1 EEFLLPKKLCQVLDHVISNDVCAOV 25

RESULT 6  
 ID W13649 standard; Protein: 261 AA.  
 AC W13649.  
 DT 24-FEB-1998 (first entry)  
 DE Human prostatic specific antigen.  
 KW Human; prostatic specific antigen; probasin; rat; cancer;  
 OS Homo sapiens.  
 PN W09711172-AL.  
 PD 27-MAR-1997.  
 PF 20-SEP-1995; U15123.  
 PR 20-SEP-1995; US-004044.  
 PA (MORC-) WORCESTER FOUND BIOMEDICAL RES.  
 PI zamecnik PA;  
 DR WPI: 97-202881/18.  
 DR N-PSDB: 761701.  
 PT Treatment of benign prostatic hyperplasia or prostatic cancer  
 PT using an anti-sense oligonucleotide targeted against prostate  
 PT specific antigen or the probasin gene  
 PS Claim 2; Page 23-30; 49pp; English.  
 CC A novel method has been developed for treating a patient diagnosed as  
 CC having benign prostatic hyperplasia or a prostatic cancer. The method  
 CC comprises administering to the patient a therapeutic amount of a  
 CC composition comprising an antisense oligonucleotide which selectively  
 CC hybridizes to prostate specific antigen (PSA) gene or mRNA or to a  
 CC probasin gene or mRNA sequence of the patient, where the antisense  
 CC oligonucleotide inhibits expression of the sequence. The present  
 CC sequence represents human PSA. The antisense oligonucleotide can be  
 CC used to effectively inhibit the growth of (and for the killing of)  
 CC of the PSA and probasin genes is specific to prostatic origin. Expression  
 CC antisense oligonucleotide can be administered systemically, making it  
 CC particularly useful in late stage prostatic cancer which has  
 CC metastasized, and in which the cells have become resistant to oestrogen  
 CC or anti-androgen therapy. The method can also be used in benign prostate  
 CC hyperplasia, or early stage prostatic cancer, as an alternative for the  
 CC more radical procedures currently used, such as transurethral resection,  
 CC radical prostatectomy, or physical or chemical castration.  
 SQ Sequence 261 AA;

Query Match 100.0%; Score 183; DB 1; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 3,20e-13;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 163 EEFLLPKKLCQVLDHVISNDVCAOV 187

QY 1 EEFLLPKKLCQVLDHVISNDVCAOV 25

RESULT 7  
 ID R84671 standard; Protein: 237 AA.  
 AC R84671;  
 DT 04-MAR-1996 (first entry)  
 DE Mature kallikrein hK3.  
 KW Kallikrein hK3; serine protease; prostate-specific antigen;  
 KW prostate cancer; diagnosis.  
 OS Homo sapiens.  
 PN W0950758-AL.  
 PD 16-NOV-1995.  
 PF 09-MAY-1995; U06157.  
 PR 10-MAY-1994; US-241174.  
 PR 02-MAY-1995; US-42767.  
 PA (HYBR-) HYBRTECH INC.  
 PA (MAYO-) MAYO FOUNDATION.  
 PI Saedi MS, Tindall DJ, Young CYF;  
 DR WPI: 95-404123/51.  
 PT New isolated prostate-specific kallikrein polypeptide(s) - used to  
 PT develop prods. for use in assays for such polypeptide(s), partic.  
 PT for diagnosis and monitoring of prostate cancer  
 PS Disclosure: Page 31; 61pp; English.  
 CC Human kallikrein hK3 (R84671) is a prostate-specific kallikrein  
 CC showing homology to prostate-specific glandular kallikrein hK2  
 CC (R84667). Nonhomologous regions of the kallikreins can be used  
 CC for prep. of antibodies specific to hK2.  
 SQ Sequence 237 AA;

Query Match 97.3%; Score 178; DB 1; Length 237;  
 Best Local Similarity 96.0%; Pred. No. 1.34e-12;  
 Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 139 EEFLLPKKLCQVLDHVISNDVCAOV 163  
 1 EEFLLPKKLCQVLDHVISNDVCAOV 25

RESULT 8  
 ID W45398 standard; Protein: 237 AA.  
 AC W45398;  
 DT 06-JUL-1998 (first entry)  
 DE Prostate-specific antigen protein hK3 (PSA).  
 KW Prostate-specific antigen; hK3; PSA; human; hK2;  
 KW prostate-specific glandular kallikrein; prostate carcinoma;  
 KW prostate cancer; benign prostatic hyperplasia; diagnosis.  
 OS Homo sapiens.  
 PN W09802748-AL.  
 PD 22-JAN-1998.  
 PF 15-JUL-1997; U12322.  
 PR 15-JUL-1996; US-680868.  
 PA (HYBR-) HYBRTECH INC.  
 PA (MAYO-) MAYO FOUNDATION.  
 PI Grauer U, Klee GG, Mikolajczyk SD, Saedi M, Tindall DJ,  
 PI Young CYF;  
 DR WPI: 98-120378/11.  
 PT Diagnostic methods using antibodies which bind prostate antigens -  
 PT useful for e.g. monitoring treatment or progression of prostate  
 PT cancer  
 PS Disclosure: Page 55-56; 100pp; English.  
 CC This polypeptide comprises human prostate-specific antigen (PSA)  
 CC hK3. PSA levels are widely used as a prognostic indicator of  
 CC prostate carcinoma. The invention provides a novel diagnostic  
 CC method comprising contacting antibodies that specifically bind to  
 CC human prostate-specific glandular kallikrein protein hK2 (see  
 CC W45395) or pro-hK2 (hK2, see W45396), but not with PSA, with a  
 CC sample of physiological fluid from a human. The assay is based on  
 CC the discovery that hK2 is detected in the supernatant of a  
 CC prostate carcinoma cell line and that hK2 is present in human  
 CC physiological fluid from prostate cancer cells. The method is  
 CC useful for monitoring the treatment and/or progression of prostate  
 CC cancer, or for the early detection of prostate cancer in males that

CC have BPH or a high grade prostatic neoplasia (HPGN) or whose family  
CC members have or had BPH, HPGN or a prostate cancer. Monitoring the  
CC presence and/or amount of hK2 complexes with plasma proteins may  
CC also be important in distinguishing between prostate cancer and BPH.  
SQ Sequence 237 AA:

Query Match 97.3%; Score 178; DB 1; Length 237;  
Best Local Similarity 96.0%; Pred. No. 1.34e-12;  
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 139 EEFLEPKKLCVQLHVISNDVCAOV 163  
1 EEFLEPKKLCVQLHVISNDVCAOV 25

RESULT 9  
ID W27387 standard; peptide: 30 AA.

AC W27387;  
DT 14-APR-1998 (first entry)  
DE Prostate specific antigen oligo-epitope peptide.  
KW Prostate specific antigen; PSA; epitope; vaccine; immune response;  
KW Prostatic tumor cell; human leukocyte antigen; HLA; cytotoxic;  
KW T lymphocyte; cancer.  
OS Synthetic.  
OS Homo sapiens.  
PN MO9735021-AZ.  
PD 25-SEP-1997.  
PE 19-MAR-1997; U04454.  
PR 20-MAR-1996; US-618936.  
PA (USSH) US SEC DEPT HEALTH.  
PI Schlom J, Tsang K, Zaremba S;  
DR N-PSDB: T90073.  
DR Prostate specific antigen multiple epitope peptide - useful in  
PT vaccine to produce immune response against prostate specific antigen  
PT in prostatic tumor cell  
PS Claim 13; Page 63; 85pp; English.  
CC The present sequence represents a prostate specific antigen  
CC oligo-epitope peptide (PSA-OP). The PSA-OP comprises more than 1  
CC adjoined PSA-OP which generates a prostate specific response in  
CC a portion of the human population. The PSA-OP can be used to produce  
CC PSA specific cytotoxic T lymphocytes, in vivo or in vitro which can  
CC kill or prevent the establishment or growth of PSA positive tumor  
CC cells, i.e. to prevent or treat prostatic cancer. The products may be  
CC used to elicit a PSA specific immune response. When a recombinant virus  
CC is used to elicit the response it is coadministered with a carrier, a  
CC biological response modifier e.g. Interleukin 2, 6 or 12, interferon,  
CC tumour necrosis factor, granulocyte/macrophage-colony stimulating factor  
CC or cyclophosphamide and an adjuvant selected from RIBI Detox, QS21,  
CC alum and incomplete Freund's adjuvant. PSA-OP comprises multiple PSA  
CC epitopes, allowing an immune response to PSA in a population of humans  
CC having more than 1 HLA Class I molecule type to be generated.

SQ Sequence 30 AA:

Query Match 93.4%; Score 171; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. No. 9.87e-12;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 FLTPKRLQCVLDHVISNDVCAOV 23  
3 FLTPKRLQCVLDHVISNDVCAOV 25

RESULT 10  
ID W11023 standard; Protein: 240 AA.

AC W11023;  
DT 27-OCT-1997 (first entry)  
DE Human prostate specific antigen.  
KW PSA; kallikrein; prostate cancer; antigenic; polyclonal antisera.  
OS Homo sapiens.  
PN WO9640754-A1.  
PD 19-DEC-1996.

PF 06-JUN-1996; U09303.  
PR 07-JUN-1995; US-472228.  
PA (TEXA) UNIV TEXAS SYSTEM.  
PI Fritsche HA, Johnston DA, Kokolus WJ;  
DR WPI: 97-108633/10.  
PT Antigen peptide derived from prostate specific antigen - does not  
PT cross react with related kallikreins, for diagnosis of prostate  
PT cancer.  
PS Claim 1; Fig 1; 74pp; English.  
CC The present sequence represents human prostate specific antigen (PSA)  
CC which is used to produce novel peptides derived from the 240 residue  
CC sequence. The preferred peptides have two hydrophobic regions and one  
CC hydrophilic region each of about 5 aa in length, arranged as follows:  
CC hydrophobic-hydrophilic-hydrophobic. Polyclonal antisera specific for  
CC the antigenic peptides may be used in a method for diagnosing prostate  
CC cancer in vivo or in vitro. The peptides represent a sequence unique  
CC to PSA which does not cross react with certain kallikreins.  
SQ Sequence 240 AA:

Query Match 73.2%; Score 134; DB 1; Length 240;  
Best Local Similarity 94.7%; Pred. No. 3.20e-07;  
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 148 KKLQCVLDHVISNDVCAOV 166  
7 KKLQCVLDHVISNDVCAOV 25

RESULT 11  
ID W96186 standard; Protein: 237 AA.

AC W96186;  
DT 27-APR-1999 (first entry)  
DE Mature human kallikrein 2 (hk2).  
KW Human kallikrein 2; hk2; breast cancer; diagnosis; detection;  
KW treatment; monitoring; prostate specific antigen; PSA.  
OS Homo sapiens.  
PN WO9659073-A1.  
PD 30-DEC-1998.  
PE 19-JUN-1998; U12840.  
PR 20-JUN-1997; US-050963.  
PA (KLEE) KLEE G G.  
PA (MAYO) MAYO FOUNDATION.  
PA (TIND) TINDALL D J.  
PA (YOUNG) YOUNG C Y F.  
PI Klee GG, Tindall DJ, Young CYF;  
DR WPI: 99-105632/09.  
DR N-PSDB: X08946.

PT Use of human kallikrein 2 - as a marker for developing products for  
PT the diagnosis, prognosis, monitoring and treatment of breast cancer  
PT disclosure; Page 31; 67pp; English.  
PS Human kallikrein 2 (hk2) is expressed at elevated levels relative to  
CC the prostate cancer antigen, prostate specific antigen (PSA) by  
CC breast cancer cells when stimulated by androgens. Detecting levels  
CC of hk2 may provide a simple diagnostic tool for detecting or  
CC determining breast cancer. Detecting hk2 is achieved by producing  
CC an hk2 DNA from hk2 RNA by reverse transcription. The hk2 DNA  
CC obtained is then amplified by PCR and detected using probes.  
CC Determination of breast cancer may also be determined by exposing  
CC the hk2 polypeptide to an agent which binds to it and then detecting  
CC the binary complex formed. The amount of complex formed correlates  
CC directly to the presence of breast cancer in the human individual.  
CC The methods can be used in the diagnosis, treatment and/or  
CC monitoring of the progression or remission of breast cancer and/or  
CC monitoring hk2 levels.  
SQ Sequence 237 AA:

Query Match 71.0%; Score 130; DB 1; Length 237;  
Best Local Similarity 66.7%; Pred. No. 9.62e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 139 EEFLEPKKLCVQLHVISNDVCAOV 162  
1 EEFLEPKKLCVQLHVISNDVCAOV 24



KM Prostate cancer: detection: hk2; hk1; hk3; phk2; tissue kallikrein;  
 KM pphk2; prostate-specific glandular kallikrein protein; PSA; human;  
 KW prostate-specific antigen.  
 OS Homo sapiens.  
 OS Synthetic.  
 PN MO9846795-A1.  
 PD 22-OCT-1998.  
 PF 09-APR-1998: 007027.  
 PR 11-APR-1997; US-843076.  
 PA (BAYO ) BAYLOR COLLEGE MEDICINE.  
 PA (MAYO-) MAYO FOUNDATION.  
 PI Slawin KM, Tindall DJ, Young CYF;  
 DR WPI: 98-594592/50.  
 PT Detection of human kallikrein 2 RNA - by reverse transcription and  
 PT amplification by PCR, for detecting, monitoring and staging of  
 PT prostate cancer.  
 PS Disclosure: Page 70-71; 90pp; English.  
 CC The present invention describes a diagnostic method for detecting human  
 CC kallikrein 2 (hk2) DNA. The method comprises: (a) contacting DNA obtained  
 CC by reverse transcription (RT) of RNA from a human physiological sample  
 CC which comprises cells suspected of containing hk2 RNA with at least 2  
 CC oligonucleotides to amplify the DNA by PCR to yield amplified hk2 DNA,  
 CC where the conditions amplify the DNA obtained by RT of RNA from at least  
 CC one cell containing hk2 in a sample which comprises at least 107 to 109  
 CC cells; and (b) detecting the presence of the amplified hk2 DNA. The  
 CC method can be used for detecting, monitoring the progression of and  
 CC pathologically staging prostate cancer. The present sequence represents  
 CC a hk2 variant.  
 SQ Sequence 237 AA;

Query Match 71.0%; Score 130; DB 1; Length 237;  
 Best Local Similarity 66.7%; Pred. No. 9.62e-07;  
 Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 139 EEFLRPSIOCVSLHLISNDWCAR 162  
 |||||:|||||::|||:|  
 QY 1 EEFLPKRLQCVDLHVISNDVCAQ 24

Search completed: Thu Jun 22 21:23:30 2000  
 Job time : 8 secs.

# WIRE

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MSRCH\_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 22 21:23:47 2000; MasPar time 3.35 Seconds

Tabular output not generated. 107.731 Million cell updates/sec

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 183  
Sequence: 1 EEFLEPKLQCVLDHVISNDVCAQV 25

Scoring table:  
PAM 150  
Gap 15

Searched: 145341 segs, 14437480 residues

Post-processing: Minimum Match 08  
Listing first 45 summaries

Database: a-issued  
1:5A\_COMB 2:5B\_COMB 3:6\_COMB 4:PCT\_COMB 5:backfiles1

Statistics: Mean 19.756; Variance 64.432; scale 0.307

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	183	100.0	237	2	US-08-844-	Sequence 2, Applicatio	3.60e-13
2	183	100.0	237	2	US-08-718-	Sequence 2, Applicatio	3.60e-13
3	183	100.0	237	3	US-08-768-	Sequence 1, Applicatio	3.60e-13
4	183	100.0	261	1	US-08-744-	Sequence 5, Applicatio	3.60e-13
5	183	100.0	261	2	US-09-102-	Sequence 5, Applicatio	3.60e-13
6	178	97.3	237	4	PCT-US94-0	Sequence 11, Applicati	1.45e-12
7	178	97.3	237	1	US-08-096-	Sequence 11, Applicati	1.45e-12
8	178	97.3	237	1	PCT-US95-0	Sequence 1, Applicatio	2.58e-07
9	178	97.3	237	1	US-08-472-	Sequence 1, Applicatio	2.58e-07
10	178	97.3	237	1	US-08-096-	Sequence 1, Applicatio	2.58e-07
11	178	97.3	237	1	US-08-472-	Sequence 1, Applicatio	2.58e-07
12	178	97.3	237	1	US-08-096-	Sequence 1, Applicatio	2.58e-07
13	178	97.3	237	1	US-08-472-	Sequence 1, Applicatio	2.58e-07
14	178	97.3	237	1	US-08-096-	Sequence 1, Applicatio	2.58e-07
15	178	97.3	237	1	US-08-472-	Sequence 1, Applicatio	2.58e-07
16	178	97.3	237	1	US-08-096-	Sequence 1, Applicatio	2.58e-07
17	178	97.3	237	1	US-08-472-	Sequence 1, Applicatio	2.58e-07
18	178	97.3	237	1	US-08-096-	Sequence 1, Applicatio	2.58e-07
19	178	97.3	237	1	US-08-472-	Sequence 1, Applicatio	2.58e-07
20	178	97.3	237	1	US-08-096-	Sequence 1, Applicatio	2.58e-07
21	178	97.3	237	1	US-08-472-	Sequence 1, Applicatio	2.58e-07
22	178	97.3	237	1	US-08-096-	Sequence 1, Applicatio	2.58e-07
23	178	97.3	237	1	US-08-472-	Sequence 1, Applicatio	2.58e-07

24	98	53.6	18	1	US-08-096-	Sequence 4, Applicatio	3.47e-03
25	96	52.5	20	1	US-08-472-	Sequence 22, Applicati	5.79e-03
26	96	52.5	20	4	PCT-US96-0	Sequence 22, Applicati	5.79e-03
27	95	51.9	258	1	US-08-744-	Sequence 3, Applicatio	7.47e-03
28	95	51.9	258	2	US-09-103-	Sequence 3, Applicatio	7.47e-03
29	95	49.7	262	2	US-08-824-	Sequence 3, Applicatio	2.07e-02
30	91	49.7	262	2	US-08-681-	Sequence 4, Applicatio	2.07e-02
31	91	49.7	262	2	US-08-790-	Sequence 4, Applicatio	2.07e-02
32	91	49.7	262	2	US-09-102-	Sequence 3, Applicatio	2.07e-02
33	91	49.7	262	2	US-08-744-	Sequence 4, Applicatio	2.07e-02
34	91	49.7	262	3	US-08-807-	Sequence 4, Applicatio	2.07e-02
35	91	49.7	262	2	US-08-790-	Sequence 1, Applicatio	2.07e-02
36	91	49.7	263	2	US-08-824-	Sequence 5, Applicatio	2.07e-02
37	91	49.7	263	3	US-08-807-	Sequence 5, Applicatio	2.07e-02
38	91	49.7	263	2	US-08-790-	Sequence 4, Applicatio	2.07e-02
39	86	47.0	232	2	US-08-897-	Sequence 31, Applicati	7.28e-02
40	86	47.0	248	1	US-09-102-	Sequence 1, Applicatio	7.28e-02
41	86	47.0	248	1	US-08-744-	Sequence 1, Applicatio	7.28e-02
42	82	44.8	228	1	US-08-472-	Sequence 7, Applicatio	1.97e-01
43	82	44.8	228	2	US-08-615-	Sequence 7, Applicatio	1.97e-01
44	82	44.8	228	1	US-08-218-	Sequence 7, Applicatio	1.97e-01
45	82	44.8	228	3	US-09-074-	Sequence 7, Applicatio	1.97e-01

## ALIGNMENTS

RESULT 1  
ID US-08-844-024-2 STANDARD; PRT: 237 AA.

Sequence 2, Application US/08844024

Sequence 2, Application US/08844024

Patent No. 5840494

GENERAL INFORMATION:

Applicant: Katz, Aaron E., et al.

TITLE OF INVENTION: A Method For Molecular Staging Of

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/844,024

FILING DATE: 18-APR-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/229,391

FILING DATE: 15-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 664-0525

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 237 amino acids

TYPE: amino acid

CC TOPOLOGY: linear  
CC MOLECULE TYPE: Protein  
SQ SEQUENCE 237 AA; 26151 MW; 312585 CN;  
Query Match 100.0%; Score 183; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3,60e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 139 EEFLTPKRLQCVDLHVISNDVCAQV 163  
1 EEFLTPKRLQCVDLHVISNDVCAQV 25  
OY  
RESULT 2  
ID US-08-718-547-2 STANDARD; PRT; 237 AA.  
AC xxxxxx  
DE Sequence 2, Application US/08718547  
CC Sequence 2, Application US/08718547  
CC Patent No. 5976794  
CC GENERAL INFORMATION:  
CC APPLICANT: Katz, Aaron E.; Buttyan, Ralph; Ratfo, Anthony;  
CC TITLE OF INVENTION: A Method For Molecular Staging Of  
CC NUMBER OF SEQUENCES: 4  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESS: Cooper & Dunham LLP  
CC STREET: 1185 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: U.S.A.  
CC ZIP: 10036  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/718,547  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: White, John P.  
CC REGISTRATION NUMBER: 28,678  
CC REFERENCE/DOCKET NUMBER: 43677-A-PCT/JPW/MUG  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 278-0400  
CC TELEFAX: (212) 391-0525  
CC TELEX:  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 237 AA; 26151 MW; 312585 CN;  
Query Match 100.0%; Score 183; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3,60e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 139 EEFLTPKRLQCVDLHVISNDVCAQV 163  
1 EEFLTPKRLQCVDLHVISNDVCAQV 25  
OY  
RESULT 3  
ID US-08-768-859A-1 STANDARD; PRT; 237 AA.

XX xxxxxx  
AC  
AC  
DE Sequence 1, Application US/08768859A  
CC Sequence 1, Application US/08768859A  
CC Patent No. 6013471  
CC GENERAL INFORMATION:  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeedi, Mohammed S.  
CC TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESS: Schwegman, Lundberg, Woessner & Kluth, P.A.  
CC STREET: P.O. Box 2938  
CC CITY: Minneapolis  
CC STATE: MN  
CC COUNTRY: USA  
CC ZIP: 55402  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/768,859A  
CC FILING DATE: 17-DECEMBER-1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Woessner, Warren D.  
CC REGISTRATION NUMBER: 30,440  
CC REFERENCE/DOCKET NUMBER: 545,002US3  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC TELEFAX: 612-339-3061  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
SQ SEQUENCE 237 AA; 26089 MW; 312729 CN;  
Query Match 100.0%; Score 183; DB 3; Length 237;  
Best Local Similarity 100.0%; Pred. No. 3,60e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 139 EEFLTPKRLQCVDLHVISNDVCAQV 163  
1 EEFLTPKRLQCVDLHVISNDVCAQV 25  
OY  
RESULT 4  
ID US-08-744-026-5 STANDARD; PRT; 261 AA.  
AC xxxxxx  
DE Sequence 5, Application US/08744026  
CC Sequence 5, Application US/08744026  
CC Patent No. 5786148  
CC GENERAL INFORMATION:  
CC APPLICANT: Bandman, Olga  
CC APPLICANT: Goli, Surya K.  
CC TITLE OF INVENTION: A NOVEL PROSTATE-SPECIFIC  
CC NUMBER OF SEQUENCES: 5  
CC





SQ	SEQUENCE	237 AA; 26066 MW; 313591 CN;
	Query Match	97.3%; Score 178; DB 1; Length 237;
	Best Local Similarity	96.0%; Pred. No. 1.45e-12;
	Matches	24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db	139 EEFLLPKKLCVQLHVISNDVCAQY 163	
Oy	1 EEFLLPKKLCVQLHVISNDVCAQY 25	
RESULT	8 STANDARD:	PRT: 237 AA.
ID	PCT-US95-06157-1	
XX	AC xxxxxx	
XX		
DT		
DE	Sequence 1, Application PC/TUS9506157	
CC	Sequence 1, Application PC/TUS9506157	
CC	GENERAL INFORMATION:	
CC	APPLICANT: Mayo Foundation for Medical Education	
CC	APPLICANT: and Research	
CC	APPLICANT: Hypritech Incorporated	
CC	APPLICANT: Tindall, Donald J.	
CC	APPLICANT: Young, Charles Y.F.	
CC	APPLICANT: Saedi, Mohammed S.	
CC	TITLE OF INVENTION: Recombinant HK2 Polypeptide	
CC	NUMBER OF SEQUENCES: 18	
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE: Schwegman, Lundberg & Woessner, P.A.	
CC	STREET: 3500 IDS Center	
CC	CITY: Minneapolis	
CC	STATE: MN	
CC	COUNTRY: USA	
CC	ZIP: 55402	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: Patentin Release #1.0, Version #1.25	
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER: PCT/US95/06157	
CC	FILING DATE:	
CC	CLASSIFICATION:	
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME: Raasch, Kevin W.	
CC	REGISTRATION NUMBER: 35,561	
CC	REFERENCE/DOCKET NUMBER: 150.148W01	
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE: 612-339-0331	
CC	TELEFAX: 612-339-3061	
CC	INFORMATION FOR SEQ ID NO: 1:	
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH: 237 amino acids	
CC	TYPE: amino acid	
CC	STRANDEDNESS: single	
CC	TOPOLOGY: linear	
CC	MOLECULE TYPE: peptide	
SQ	SEQUENCE 237 AA; 26066 MW; 313591 CN;	
	Query Match	97.3%; Score 178; DB 4; Length 237;
	Best Local Similarity	96.0%; Pred. No. 1.45e-12;
	Matches	24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db	139 EEFLLPKKLCVQLHVISNDVCAQY 163	
Oy	1 EEFLLPKKLCVQLHVISNDVCAQY 25	
RESULT	9 STANDARD:	PRT: 240 AA.
ID	PCT-US96-09303-1	



CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC IMMEDIATE SOURCE:  
CC CLONE: Amino acid sequence of HK2  
CC CLONE: (deduced from cDNA sequence)  
SQ SEQUENCE 237 AA; 26182 MW; 304329 CN;

Query Match 71.0%; Score 130; DB 1; Length 237;  
Best Local Similarity 66.7%; Pred. No. 7,58e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 139 EEFLLPRSLQCVSLHLSDMCMAR 162  
1 EEFLLPRKLCQVDLHVLSNDVCAQ 24

RESULT 12  
ID PCT-US94-07329-10 STANDARD; PRT: 237 AA.  
XX xxxxxx

Sequence 10, Application PC/TUS9407329

DE  
XX  
CC Sequence 10, Application PC/TUS9407329  
CC GENERAL INFORMATION:  
CC APPLICANT: Mayo Foundation for Medical  
CC APPLICANT: Education and Research  
CC TIME OF INVENTION: Antibodies Specific for Human  
CC TIME OF INVENTION: Prostate Glandular kallikrein  
CC NUMBER OF SEQUENCES: 11  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Schwegman, Lundberg & Woessner  
CC STREET: 3500 IDS Center  
CC STREET: 80 South Eighth Street  
CC CITY: Minneapolis  
CC STATE: MN  
CC COUNTRY: USA  
CC ZIP: 55402  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US94/07329

CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Raasch, Kevin W.  
CC REGISTRATION NUMBER: 35,651  
CC REFERENCE/DOCKET NUMBER: 150.62WO  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC INFORMATION FOR SEQ ID NO: 10:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC IMMEDIATE SOURCE:  
CC CLONE: Amino acid sequence of HK2  
CC CLONE: (deduced from cDNA sequence)  
SQ SEQUENCE 237 AA; 26182 MW; 304329 CN;

Query Match 71.0%; Score 130; DB 4; Length 237;  
Best Local Similarity 66.7%; Pred. No. 7,58e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 139 EEFLLPRSLQCVSLHLSDMCMAR 162  
1 EEFLLPRKLCQVDLHVLSNDVCAQ 24

RESULT 13  
ID PCT-US95-06157-16 STANDARD; PRT: 237 AA.  
XX xxxxxx

Sequence 16, Application PC/TUS9506157

DE  
XX  
CC Sequence 16, Application PC/TUS9506157  
CC GENERAL INFORMATION:  
CC APPLICANT: Mayo Foundation for Medical Education  
CC APPLICANT: and Research  
CC APPLICANT: Hydritech Incorporated  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeedi, Mohammed S.  
CC TITLE OF INVENTION: Recombinant HK2 Polypeptide  
CC NUMBER OF SEQUENCES: 18  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Schwegman, Lundberg & Woessner, P.A.  
CC STREET: 3500 IDS Center  
CC CITY: Minneapolis  
CC STATE: MN  
CC COUNTRY: USA  
CC ZIP: 55402  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/06157

CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Raasch, Kevin W.  
CC REGISTRATION NUMBER: 35,561  
CC REFERENCE/DOCKET NUMBER: 150.148WO1  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC INFORMATION FOR SEQ ID NO: 16:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
SQ SEQUENCE 237 AA; 26182 MW; 304329 CN;

Query Match 71.0%; Score 130; DB 4; Length 237;  
Best Local Similarity 66.7%; Pred. No. 7,58e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 139 EEFLLPRSLQCVSLHLSDMCMAR 162  
1 EEFLLPRKLCQVDLHVLSNDVCAQ 24

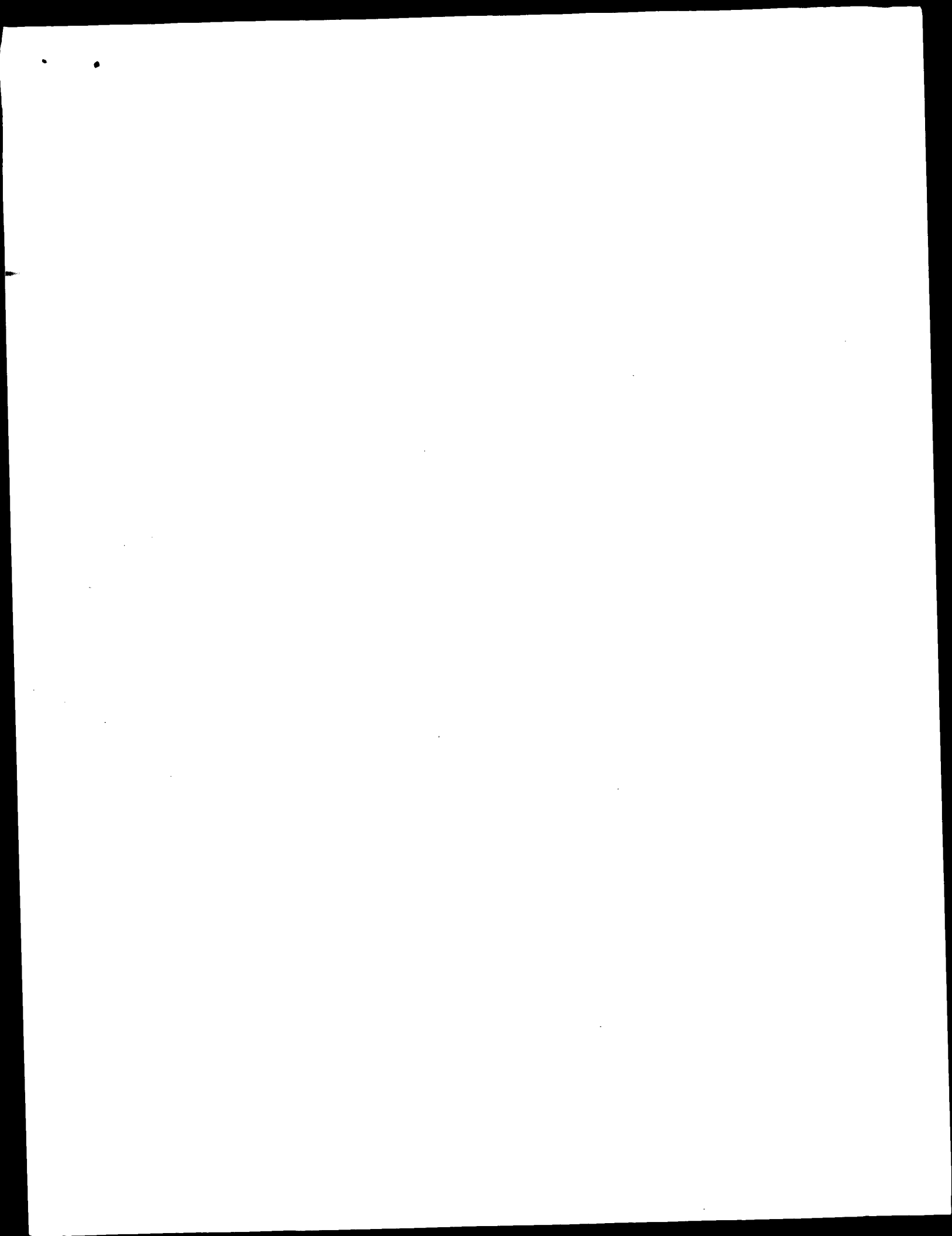
RESULT 14  
ID US-08-768-859A-21 STANDARD; PRT: 237 AA.  
XX xxxxxx

DT  
XX

DE Sequence 21, Application US/08768859A  
XX  
CC Sequence 21, Application US/08768859A  
CC Patent No. 6013471  
CC GENERAL INFORMATION:  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeed, Mohammed S.  
CC TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESSES:  
CC ADDRESS: Schwegman, Lundberg, Woessner & Kluth, P.A.  
CC STREET: P.O. Box 2938  
CC CITY: Minneapolis  
CC STATE: MN  
CC COUNTRY: USA  
CC ZIP: 55402  
CC  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/768,859A  
CC FILING DATE: 17-DECEMBER-1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Woessner, Warren D.  
CC REGISTRATION NUMBER: 30,440  
CC REFERENCE/DOCKET NUMBER: 545.002US3  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC TELEFAX: 612-339-3061  
CC INFORMATION FOR SEQ ID NO: 21:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC SEQUENCE 237 AA: 26187 MW: 308471 CN;  
SO  
Query Match 71.0%; Score 130; DB 3; Length 237;  
Best Local Similarity 66.7%; Pred. No. 7.58e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
DB 139 EEFLLPRSLQCVSLHLSNDMCA 162  
QY 1 EEFLLPRSLQCVSLHLSNDVCAQ 24  
RESULT 15  
ID US-08-768-859A-16 STANDARD: PRT: 237 AA.  
AC xxxxxx  
DT  
XX  
DE Sequence 16, Application US/08768859A  
CC  
CC Patent No. 6013471  
CC GENERAL INFORMATION:  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeed, Mohammed S.  
CC TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESSES:  
CC ADDRESS: Schwegman, Lundberg, Woessner & Kluth, P.A.  
CC STREET: P.O. Box 2938  
CC CITY: Minneapolis  
CC STATE: MN

CC COUNTRY: USA  
CC ZIP: 55402  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/768,859A  
CC FILING DATE: 17-DECEMBER-1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Woessner, Warren D.  
CC REGISTRATION NUMBER: 30,440  
CC REFERENCE/DOCKET NUMBER: 545.002US3  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC TELEFAX: 612-339-3061  
CC INFORMATION FOR SEQ ID NO: 16:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 237 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC SEQUENCE 237 AA: 26139 MW: 304348 CN;  
SO  
Query Match 71.0%; Score 130; DB 3; Length 237;  
Best Local Similarity 66.7%; Pred. No. 7.58e-07;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
DB 139 EEFLLPRSLQCVSLHLSNDMCA 162  
QY 1 EEFLLPRSLQCVSLHLSNDVCAQ 24

Search completed: Thu Jun 22 21:23:55 2000  
Job time : 8 secs.



TRANSLATION AA-&gt;NA

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mparch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 15:48:17 2000; Maspar time 285.75 Seconds

Tabular output not generated. 1118.717 Million cell updates/sec

Title: >US-09-332-866-1  
Description: (123) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYVTNACNCNNA.....AYGAYTNGYGCNCARGTN 75  
Comp: CTRYTAARANTWGNGNTT.....TTCCTCANACACGNGNTYCAN

Scoring table: TABLE bkttranslate2  
Gap 40

Match STD : Dbase 0; Query 0

Searched: 23132 seqs, 213114489 bases x 2

Post-processing: Minimum Match 08

Listing first 45 summaries

Database:

emb161  
1:em\_ba1 2:em\_ba2 3:em\_fun 4:em\_htg1 5:em\_htg2 6:em\_htg3  
7:em\_hum1 8:em\_hum2 9:em\_hum3 10:em\_hum4 11:em\_in  
12:em\_com 13:em\_ov 14:em\_ov 15:em\_pat 16:em\_pi 17:em\_ro  
18:em\_sts 19:em\_un 20:em\_v1  
genbank116  
21:gb\_htg1 22:gb\_htg10 23:gb\_htg11 24:gb\_htg12  
25:gb\_htg13 26:gb\_htg14 27:gb\_htg2 28:gb\_htg3 29:gb\_htg4  
30:gb\_htg5 31:gb\_htg6 32:gb\_htg7 33:gb\_htg8 34:gb\_htg9

Database:

Statistics: Mean 78.372; Variance 327.890; scale 0.239

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the total being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description	Pred. No.
1	277	87.7	40334	28	AC011523	Homo sapiens chromosome 7.32e+08			
2	145	45.9	40871	31	AC020019	Drosophila melanogaster 5.49e+01			
3	137	43.4	2286	31	AC019591	Drosophila melanogaster 1.43e+02			
4	137	43.4	60074	29	AC013501	Homo sapiens chromosome 1.43e+02			
5	137	43.4	174801	29	AC014376	Drosophila melanogaster 1.43e+02			
6	137	43.4	200319	27	AL135931	Homo sapiens chromosome 1.43e+02			
7	137	43.4	135499	24	AC013342	Homo sapiens chromosome 1.43e+02			
8	135	42.7	135499	24	AC008969	Homo sapiens chromosome 1.43e+02			
9	132	41.8	17438	29	AC014821	Drosophila melanogaster 2.53e+02			
10	132	41.8	148162	33	AC012865	Homo sapiens chromosome 2.53e+02			
11	132	41.8	149408	31	AC013582	Homo sapiens chromosome 2.53e+02			
12	132	41.8	151670	28	AC007777	Homo sapiens chromosome 2.53e+02			

13	132	41.8	156492	32	AC008460	Homo sapiens chromosome 2.53e+02			
14	132	41.8	182606	27	AL136986	Homo sapiens chromosome 2.53e+02			
15	132	41.8	202335	33	AC020603	Homo sapiens clone RPI 2.53e+02			
16	131	41.5	48836	31	AC020151	Drosophila melanogaster 2.83e+02			
17	131	41.5	112958	29	AC010107	Drosophila melanogaster 2.83e+02			
18	131	41.5	129093	31	AC011483	Homo sapiens chromosome 2.83e+02			
19	130	41.1	166880	28	AC009756	Homo sapiens chromosome 2.83e+02			
20	130	41.1	276261	21	HSAC000406	Homo sapiens clone 45-3.17e+02			
21	127	40.2	10484	29	AC014515	Homo sapiens chromosome 4.39e+02			
22	127	40.2	17438	29	AC014821	Drosophila melanogaster 4.39e+02			
23	127	40.2	119844	24	AC010044	Drosophila melanogaster 4.39e+02			
24	127	40.2	121576	24	AC010043	Drosophila melanogaster 4.39e+02			
25	127	40.2	143126	32	AC012603	Homo sapiens chromosome 4.39e+02			
26	127	40.2	148476	24	AC020751	Homo sapiens chromosome 4.39e+02			
27	127	40.2	162985	34	AC020751	Homo sapiens chromosome 4.39e+02			
28	127	40.2	166290	30	AC010815	Homo sapiens clone RPI 4.39e+02			
29	127	40.2	223724	21	AP000834	Homo sapiens chromosome 4.39e+02			
30	125	39.6	283187	31	AC010615	Homo sapiens chromosome 5.44e+02			
31	125	39.6	71307	28	AC008296	Drosophila melanogaster 5.44e+02			
32	125	39.6	83714	33	AC008297	Drosophila melanogaster 5.44e+02			
33	125	39.6	97308	29	AC013925	Drosophila melanogaster 5.44e+02			
34	125	39.6	131199	23	AC007807	Drosophila melanogaster 5.44e+02			
35	125	39.6	132819	28	AC009335	Homo sapiens chromosome 5.44e+02			
36	125	39.6	188432	34	AC018497	Homo sapiens chromosome 5.44e+02			
37	122	38.6	86367	29	AC013310	Homo sapiens chromosome 7.45e+02			
38	122	38.6	95750	34	AC020580	Homo sapiens clone RPI 7.45e+02			
39	122	38.6	126429	31	AC010447	Homo sapiens chromosome 7.45e+02			
40	122	38.6	145619	31	AC011059	Homo sapiens chromosome 7.45e+02			
41	122	38.6	146462	31	AC015474	Homo sapiens clone RPI 7.45e+02			
42	122	38.6	173737	33	AC022245	Homo sapiens chromosome 7.45e+02			
43	122	38.6	175314	31	AC021653	Homo sapiens chromosome 7.45e+02			
44	122	38.6	216858	32	AC009060	Homo sapiens chromosome 7.45e+02			
45	122	38.6	228143	31	AC010422	Homo sapiens chromosome 7.45e+02			

## ALIGNMENTS

RESULT	1	AC011523	40334 bp	DNA	HTG	07-OCT-1999
LOCUS		HOMO SAPIENS	Chromosome 19	clone LINT-F_214C7	*** SEQUENCING IN	
DEFINITION		AC011523	***, 1 ordered pieces.			
ACCESSION		AC011523.1	GI:6015229			
VERSION		HTG: HTGS_PHASE2.				
KEYWORDS		human.				
SOURCE		Homo sapiens				
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
REFERENCE		Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS		1 (bases 1 to 40334)				
TITLE		DOE Joint Genome Institute.				
JOURNAL		Sequencing of Human Chromosome 19				
REFERENCE		Unpublished				
AUTHORS		2 (bases 1 to 40334)				
TITLE		DOE Joint Genome Institute.				
JOURNAL		Direct Submission				
COMMENT		Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
		www.jgi.doe.gov#mapping database ID 25479.				
		* NOTE: This is a "working draft" sequence. It currently				
		* consists of 1 contigs. Gaps between the contigs				
		* are represented as runs of N. The order of the pieces				
		* is believed to be correct as given, however the sizes				
		* of the gaps between them are based on estimates that have				
		* provided by the submitter.				
		* This sequence will be replaced				
		* by the finished sequence as soon as it is available and				
		* the accession number will be preserved.				
		1 40334: contig of 40334 bp in length.				
		Location/Qualifiers				
		1. 40334				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/chromosome="19"				

## FEATURES

source





```

* 1892 2846: contig of 955 bp in length
* 2847 3805: contig of 959 bp in length
* 3806 4732: contig of 927 bp in length
* 4733 5763: contig of 1031 bp in length
* 5764 6689: contig of 926 bp in length
* 6690 7658: contig of 969 bp in length
* 7659 8587: contig of 929 bp in length
* 8588 9528: contig of 941 bp in length
* 9529 10461: contig of 933 bp in length
* 10462 11355: contig of 894 bp in length
* 11356 12275: contig of 920 bp in length
* 12276 13237: contig of 962 bp in length
* 13238 14152: contig of 915 bp in length
* 14153 15096: contig of 944 bp in length
* 15097 16044: contig of 948 bp in length
* 16045 16968: contig of 924 bp in length
* 16969 17899: contig of 931 bp in length
* 17900 18769: contig of 870 bp in length
* 18770 19691: contig of 922 bp in length
* 19692 20665: contig of 974 bp in length
* 20666 21584: contig of 919 bp in length
* 21585 22514: contig of 930 bp in length
* 22515 23431: contig of 917 bp in length
* 23432 24389: contig of 958 bp in length
* 24390 25310: contig of 921 bp in length
* 25311 26249: contig of 939 bp in length
* 26250 27186: contig of 937 bp in length
* 27187 28104: contig of 918 bp in length
* 28105 29003: contig of 899 bp in length
* 29004 29914: contig of 911 bp in length
* 29915 30837: contig of 923 bp in length
* 30838 31793: contig of 956 bp in length
* 31794 32742: contig of 949 bp in length
* 32743 33700: contig of 958 bp in length
* 33701 34645: contig of 945 bp in length
* 34646 35598: contig of 953 bp in length
* 35599 36521: contig of 923 bp in length

```

```

* 36522 37474: gap of unknown length
* 37475 38421: gap of unknown length
* 38422 39375: gap of unknown length
* 39376 40312: gap of unknown length
* 40313 41271: gap of unknown length
* 41272 42252: gap of unknown length
* 42253 43176: gap of unknown length
* 43177 44128: gap of unknown length
* 44129 45093: gap of unknown length
* 45094 46032: gap of unknown length
* 46033 46967: gap of unknown length
* 46968 47892: gap of unknown length
* 47893 48847: gap of unknown length
* 48848 49774: gap of unknown length
* 49775 50705: gap of unknown length
* 50706 51624: gap of unknown length
* 51625 52536: gap of unknown length
* 52537 53469: gap of unknown length
* 53470 54421: gap of unknown length
* 54422 55350: gap of unknown length
* 55351 56288: gap of unknown length
* 56289 57228: gap of unknown length
* 57229 58152: gap of unknown length
* 58153 59155: gap of unknown length
* 59156 60074: gap of unknown length
* 60074 60074: Location/Qualifiers
  1. 60074
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="3"
    /map="3"
    /clone="219_A_11"
    /clone_id="RPCT-11 Human Male BAC"
BASE COUNT 15665 a 13176 c 15660 g 15033 t 540 others
ORIGIN
Query Match 43.4% Score 137: DB 29: Length 60074:
Best Local Similarity 37.7% Pred. No. 1.43e+02:
Matches 26: Conservative 18: Mismatches 25: Indels 0: Gaps 0:
DB 50397 AGTCTTTCATCCAGAGCTCTTATCATGATTTGCCGGCCATCCACATGAT 50456
OY 5 ARITYTNACNCNARARATYTCARFTGYGNGAYTYNCAYGTATWMAATGATCTNT 64
DB 50457 GTGCACAG 50465
OY 65 GYGCACAG 73

```



```

* 115310 116109: gap of 800 bp
* 116110 119236: contig of 3127 bp in length
* 119237 120036: gap of 800 bp
* 120037 123078: contig of 3042 bp in length
* 123079 123878: gap of 800 bp
* 123879 127882: contig of 4004 bp in length
* 127883 128682: gap of 800 bp
* 128683 134435: contig of 5753 bp in length
* 134436 135235: gap of 800 bp
* 135236 137655: contig of 2420 bp in length
* 137656 138455: gap of 800 bp
* 138456 140516: contig of 2061 bp in length
* 140517 141316: gap of 800 bp
* 141317 144267: contig of 2951 bp in length
* 144268 145067: gap of 800 bp
* 145068 145565: contig of 1498 bp in length
* 145566 147365: gap of 800 bp
* 147366 149673: contig of 2308 bp in length
* 149674 150473: gap of 800 bp
* 150474 152763: contig of 2290 bp in length
* 152764 153563: gap of 800 bp
* 153564 155480: contig of 1917 bp in length
* 155481 156280: gap of 800 bp
* 156281 157399: contig of 1119 bp in length
* 157400 158199: gap of 800 bp
* 158200 159896: contig of 1697 bp in length
* 159897 160696: gap of 800 bp
* 160697 161724: contig of 1028 bp in length
* 161725 162524: gap of 800 bp
* 162525 166789: contig of 4265 bp in length
* 166790 167589: gap of 800 bp
* 167590 175692: contig of 8103 bp in length
* 175693 176492: gap of 800 bp
* 176493 181011: contig of 4519 bp in length
* 181012 181811: gap of 800 bp
* 181812 184445: contig of 2634 bp in length
* 184446 185245: gap of 800 bp
* 185246 191325: contig of 6080 bp in length
* 191326 192125: gap of 800 bp
* 192126 195243: contig of 3118 bp in length
* 195244 196043: gap of 800 bp
* 196044 200319: contig of 4276 bp in length.
*
FEATURES
  Source
    1. 200319
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="1"
    /clone="RP5-932N9"
    /clone_11b="RPC1-5"
BASE COUNT  51936 a 34697 c 33442 g 49819 t 30425 others
ORIGIN
Query Match 43.4%; Score 137; DB 27; Length 200319;
Best Local Similarity 44.6%; Pred. No. 1.43e+02;
Matches 25; Conservative 13; Mismatches 18; Indels 0; Gaps 0;
DB 136893 CTTGGCCCAAAACCTCCATGTCGCTCCGATGCTTCCTTCATGTTGTTG 136948
QY 10 TTNACCCAAARATTCARTGTGTCGATGTCATGTCATGTCATGTCATGTCATG 65

```

AUTHORS Smith,D.R.  
 TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 228527)  
 AUTHORS Smith,D.R.  
 JOURNAL Direct Submission  
 Submitted (03-NOV-1999) Genome Therapeutics Corporation, 100 Beaver  
 Street, Waltham, MA 02453, USA  
 On Dec 8, 1999 this sequence version replaced gl:6456132.

```

----- Genome Center
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com
----- Project Information
Center project name: hg053
----- Summary Statistics
Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Consensus quality: 124305 bases at least Q40
Consensus quality: 169118 bases at least Q30
Consensus quality: 187476 bases at least Q20
Insert size: 228527; sum-of-contigs
Quality coverage: 2.0x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1191: contig of 1191 bp in length
1192 2264: gap of unknown length
2265 3335: contig of 1073 bp in length
3336 4495: gap of unknown length
4496 5654: gap of 1071 bp in length
5655 6842: gap of unknown length
6843 8057: gap of 1160 bp in length
8058 9158: gap of unknown length
9159 10625: gap of 1159 bp in length
10626 12209: gap of unknown length
12210 13623: gap of 1167 bp in length
13624 14631: gap of unknown length
14632 15936: gap of 1008 bp in length
15937 17353: gap of unknown length
17354 18449: gap of 1305 bp in length
18450 19983: gap of unknown length
19984 21369: gap of 1086 bp in length
21370 22537: gap of unknown length
22538 23909: gap of 1168 bp in length
23909 25338: gap of unknown length
25339 26755: gap of 1172 bp in length
26756 28173: gap of unknown length
28174 29590: gap of 1186 bp in length
29591 31007: gap of unknown length
31008 32424: gap of 1199 bp in length
32425 33841: gap of unknown length
33842 35259: gap of 1213 bp in length
35260 36677: gap of unknown length
36678 38094: gap of 1226 bp in length
38095 39511: gap of unknown length
39512 40928: gap of 1239 bp in length
40929 42346: gap of unknown length
42347 43763: gap of 1252 bp in length
43764 45180: gap of unknown length
45181 46597: gap of 1265 bp in length
46598 48014: gap of unknown length
48015 49431: gap of 1278 bp in length
49432 50849: gap of unknown length
50850 52266: gap of 1291 bp in length
52267 53683: gap of unknown length
53684 55100: gap of 1304 bp in length
55101 56517: gap of unknown length
56518 57934: gap of 1317 bp in length
57935 59352: gap of unknown length
59353 60769: gap of 1330 bp in length
60770 62187: gap of unknown length
62188 63604: gap of 1343 bp in length
63605 65021: gap of unknown length
65022 66438: gap of 1356 bp in length
66439 67855: gap of unknown length
67856 69272: gap of 1369 bp in length
69273 70689: gap of unknown length
70690 72106: gap of 1382 bp in length
72107 73523: gap of unknown length
73524 74940: gap of 1395 bp in length
74941 76357: gap of unknown length
76358 77774: gap of 1408 bp in length
77775 79191: gap of unknown length
79192 80608: gap of 1421 bp in length
80609 82025: gap of unknown length
82026 83442: gap of 1434 bp in length
83443 84859: gap of unknown length
84860 86276: gap of 1447 bp in length
86277 87693: gap of unknown length
87694 89110: gap of 1460 bp in length
89111 90527: gap of unknown length
90528 91944: gap of 1473 bp in length
91945 93361: gap of unknown length
93362 94778: gap of 1486 bp in length
94779 96195: gap of unknown length
96196 97612: gap of 1499 bp in length
97613 99029: gap of unknown length
99030 100446: gap of 1512 bp in length
100447 101863: gap of unknown length
101864 103280: gap of 1525 bp in length
103281 104697: gap of unknown length
104698 106114: gap of 1538 bp in length
106115 107531: gap of unknown length
107532 108948: gap of 1551 bp in length
108949 110365: gap of unknown length
110366 111782: gap of 1564 bp in length
111783 113199: gap of unknown length
113200 114616: gap of 1577 bp in length
114617 116033: gap of unknown length
116034 117450: gap of 1590 bp in length
117451 118867: gap of unknown length
118868 120284: gap of 1603 bp in length
120285 121701: gap of unknown length
121702 123118: gap of 1616 bp in length
123119 124535: gap of unknown length
124536 125952: gap of 1629 bp in length
125953 127369: gap of unknown length
127370 128786: gap of 1642 bp in length
128787 130203: gap of unknown length
130204 131620: gap of 1655 bp in length
131621 133037: gap of unknown length
133038 134454: gap of 1668 bp in length
134455 135871: gap of unknown length
135872 137288: gap of 1681 bp in length
137289 138705: gap of unknown length
138706 140122: gap of 1694 bp in length
140123 141539: gap of unknown length
141540 142956: gap of 1707 bp in length
142957 144373: gap of unknown length
144374 145790: gap of 1720 bp in length
145791 147207: gap of unknown length
147208 148624: gap of 1733 bp in length
148625 150041: gap of unknown length
150042 151458: gap of 1746 bp in length
151459 152875: gap of unknown length
152876 154292: gap of 1759 bp in length
154293 155709: gap of unknown length
155710 157126: gap of 1772 bp in length
157127 158543: gap of unknown length
158544 159960: gap of 1785 bp in length
159961 161377: gap of unknown length
161378 162794: gap of 1798 bp in length
162795 164211: gap of unknown length
164212 165628: gap of 1811 bp in length
165629 167045: gap of unknown length
167046 168462: gap of 1824 bp in length
168463 169879: gap of unknown length
169880 171296: gap of 1837 bp in length
171297 172713: gap of unknown length
172714 174130: gap of 1850 bp in length
174131 175547: gap of unknown length
175548 176964: gap of 1863 bp in length
176965 178381: gap of unknown length
178382 179798: gap of 1876 bp in length
179799 181215: gap of unknown length
181216 182632: gap of 1889 bp in length
182633 184049: gap of unknown length
184050 185466: gap of 1902 bp in length
185467 186883: gap of unknown length
186884 188300: gap of 1915 bp in length
188301 189717: gap of unknown length
189718 191134: gap of 1928 bp in length
191135 192551: gap of unknown length
192552 193968: gap of 1941 bp in length
193969 195385: gap of unknown length
195386 196802: gap of 1954 bp in length
196803 198219: gap of unknown length
198220 199636: gap of 1967 bp in length
199637 201053: gap of unknown length
201054 202470: gap of 1980 bp in length
202471 203887: gap of unknown length
203888 205304: gap of 1993 bp in length
205305 206721: gap of unknown length
206722 208138: gap of 2006 bp in length
208139 209555: gap of unknown length
209556 210972: gap of 2019 bp in length
210973 212389: gap of unknown length
212390 213806: gap of 2032 bp in length
213807 215223: gap of unknown length
215224 216640: gap of 2045 bp in length
216641 218057: gap of unknown length
218058 219474: gap of 2058 bp in length
219475 220891: gap of unknown length
220892 222308: gap of 2071 bp in length
222309 223725: gap of unknown length
223726 225142: gap of 2084 bp in length
225143 226559: gap of unknown length
226560 227976: gap of 2097 bp in length
227977 229393: gap of unknown length
229394 230810: gap of 2110 bp in length
230811 232227: gap of unknown length
232228 233644: gap of 2123 bp in length
233645 235061: gap of unknown length
235062 236478: gap of 2136 bp in length
236479 237895: gap of unknown length
237896 239312: gap of 2149 bp in length
239313 240729: gap of unknown length
240730 242146: gap of 2162 bp in length
242147 243563: gap of unknown length
243564 244980: gap of 2175 bp in length
244981 246397: gap of unknown length
246398 247814: gap of 2188 bp in length
247815 249231: gap of unknown length
249232 250648: gap of 2201 bp in length
250649 252065: gap of unknown length
252066 253482: gap of 2214 bp in length
253483 254899: gap of unknown length
254900 256316: gap of 2227 bp in length
256317 257733: gap of unknown length
257734 259150: gap of 2240 bp in length
259151 260567: gap of unknown length
260568 261984: gap of 2253 bp in length
261985 263401: gap of unknown length
263402 264818: gap of 2266 bp in length
264819 266235: gap of unknown length
266236 267652: gap of 2279 bp in length
267653 269069: gap of unknown length
269070 270486: gap of 2292 bp in length
270487 271903: gap of unknown length
271904 273320: gap of 2305 bp in length
273321 274737: gap of unknown length
274738 276154: gap of 2318 bp in length
276155 277571: gap of unknown length
277572 278988: gap of 2331 bp in length
278989 280405: gap of unknown length
280406 281822: gap of 2344 bp in length
281823 283239: gap of unknown length
283240 284656: gap of 2357 bp in length
284657 286073: gap of unknown length
286074 287490: gap of 2370 bp in length
287491 288907: gap of unknown length
288908 290324: gap of 2383 bp in length
290325 291741: gap of unknown length
291742 293158: gap of 2396 bp in length
293159 294575: gap of unknown length
294576 295992: gap of 2409 bp in length
295993 297409: gap of unknown length
297410 298826: gap of 2422 bp in length
298827 300243: gap of unknown length
300244 301660: gap of 2435 bp in length
301661 303077: gap of unknown length
303078 304494: gap of 2448 bp in length
304495 305911: gap of unknown length
305912 307328: gap of 2461 bp in length
307329 308745: gap of unknown length
308746 310162: gap of 2474 bp in length
310163 311579: gap of unknown length
311580 312996: gap of 2487 bp in length
312997 314413: gap of unknown length
314414 315830: gap of 2500 bp in length
315831 317247: gap of unknown length
317248 318664: gap of 2513 bp in length
318665 320081: gap of unknown length
320082 321498: gap of 2526 bp in length
321499 322915: gap of unknown length
322916 324332: gap of 2539 bp in length
324333 325749: gap of unknown length
325750 327166: gap of 2552 bp in length
327167 328583: gap of unknown length
328584 330000: gap of 2565 bp in length
330001 331417: gap of unknown length
331418 332834: gap of 2578 bp in length
332835 334251: gap of unknown length
334252 335668: gap of 2591 bp in length
335669 337085: gap of unknown length
337086 338502: gap of 2604 bp in length
338503 339919: gap of unknown length
339920 341336: gap of 2617 bp in length
341337 342753: gap of unknown length
342754 344170: gap of 2630 bp in length
344171 345587: gap of unknown length
345588 347004: gap of 2643 bp in length
347005 348421: gap of unknown length
348422 349838: gap of 2656 bp in length
349839 351255: gap of unknown length
351256 352672: gap of 2669 bp in length
352673 354089: gap of unknown length
354090 355506: gap of 2682 bp in length
355507 356923: gap of unknown length
356924 358340: gap of 2695 bp in length
358341 359757: gap of unknown length
359758 361174: gap of 2708 bp in length
361175 362591: gap of unknown length
362592 364008: gap of 2721 bp in length
364009 365425: gap of unknown length
365426 366842: gap of 2734 bp in length
366843 368259: gap of unknown length
368260 369676: gap of 2747 bp in length
369677 371093: gap of unknown length
371094 372510: gap of 2760 bp in length
372511 373927: gap of unknown length
373928 375344: gap of 2773 bp in length
375345 376761: gap of unknown length
376762 378178: gap of 2786 bp in length
378179 379595: gap of unknown length
379596 381012: gap of 2799 bp in length
381013 382429: gap of unknown length
382430 383846: gap of 2812 bp in length
383847 385263: gap of unknown length
385264 386680: gap of 2825 bp in length
386681 388097: gap of unknown length
388098 389514: gap of 2838 bp in length
389515 390931: gap of unknown length
390932 392348: gap of 2851 bp in length
392349 393765: gap of unknown length
393766 395182: gap of 2864 bp in length
395183 396599: gap of unknown length
396600 398016: gap of 2877 bp in length
398017 399433: gap of unknown length
399434 400850: gap of 2890 bp in length
400851 402267: gap of unknown length
402268 403684: gap of 2903 bp in length
403685 405101: gap of unknown length
405102 406518: gap of 2916 bp in length
406519 407935: gap of unknown length
407936 409352: gap of 2929 bp in length
409353 410769: gap of unknown length
410770 412186: gap of 2942 bp in length
412187 413603: gap of unknown length
413604 415020: gap of 2955 bp in length
415021 416437: gap of unknown length
416438 417854: gap of 2968 bp in length
417855 419271: gap of unknown length
419272 420688: gap of 2981 bp in length
420689 422105: gap of unknown length
422106 423522: gap of 2994 bp in length
423523 424939: gap of unknown length
424940 426356: gap of 3007 bp in length
426357 427773: gap of unknown length
427774 429190: gap of 3020 bp in length
429191 430607: gap of unknown length
430608 432024: gap of 3033 bp in length
432025 433441: gap of unknown length
433442 434858: gap of 3046 bp in length
434859 436275: gap of unknown length
436276 437692: gap of 3059 bp in length
437693 439109: gap of unknown length
439110 440526: gap of 3072 bp in length
440527 441943: gap of unknown length
441944 443360: gap of 3085 bp in length
443361 444777: gap of unknown length
444778 446194: gap of 3098 bp in length
446195 447611: gap of unknown length
447612 449028: gap of 3111 bp in length
449029 450445: gap of unknown length
450446 451862: gap of 3124 bp in length
451863 453279: gap of unknown length
453280 454696: gap of 3137 bp in length
454697 456113: gap of unknown length
456114 457530: gap of 3150 bp in length
457531 458947: gap of unknown length
458948 460364: gap of 3163 bp in length
460365 461781: gap of unknown length
461782 463198: gap of 3176 bp in length
463199 464615: gap of unknown length
464616 466032: gap of 3189 bp in length
466033 467449: gap of unknown length
467450 468866: gap of 3202 bp in length
468867 470283: gap of unknown length
470284 471700: gap of 3215 bp in length
471701 473117: gap of unknown length
473118 474534: gap of 3228 bp in length
474535 475951: gap of unknown length
475952 477368: gap of 3241 bp in length
477369 478785: gap of unknown length
478786 480202: gap of 3254 bp in length
480203 481619: gap of unknown length
481620 483036: gap of 3267 bp in length
483037 484453: gap of unknown length
484454 485870: gap of 3280 bp in length
485871 487287: gap of unknown length
487288 488704: gap of 3293 bp in length
488705 490121: gap of unknown length
490122 491538: gap of 3306 bp in length
491539 492955: gap of unknown length
492956 494372: gap of 3319 bp in length
494373 495789: gap of unknown length
495790 497206: gap of 3332 bp in length
497207 498623: gap of unknown length
498624 500040: gap of 3345 bp in length
500041 501457: gap of unknown length
501458 502874: gap of 3358 bp in length
502875 504291: gap of unknown length
504292 505708: gap of 3371 bp in length
505709 507125: gap of unknown length
507126 508542: gap of 3384 bp in length
508543 509959: gap of unknown length
509960 511376: gap of 3397 bp in length
511377 512793: gap of unknown length
512794 514210: gap of 3410 bp in length
514211 515627: gap of unknown length
515628 517044: gap of 3423 bp in length
517045 518461: gap of unknown length
518462 519878: gap of 3436 bp in length
519879 521295: gap of unknown length
521296 522712: gap of 3449 bp in length
522713 524129: gap of unknown length
524130 525546: gap of 3462 bp in length
525547 526963: gap of unknown length
526964 528380: gap of 3475 bp in length
528381 529797: gap of unknown length
529798 531214: gap of 3488 bp in length
531215 532631: gap of unknown length
532632 534048: gap of 3501 bp in length
534049 535465: gap of unknown length
535466 536882: gap of 3514 bp in length
536883 538299: gap of unknown length
538300 539716: gap of 3527 bp in length
539717 541133: gap of unknown length
541134 542550: gap of 3540 bp in length
542551 543967: gap of unknown length
543968 545384: gap of 3553 bp in length
545385 546801: gap of unknown length
546802 548218: gap of 3566 bp in length
548219 549635: gap of unknown length
549636 551052: gap of 3579 bp in length
551053 552469: gap of unknown length
552470 553886: gap of 3592 bp in length
553887 555303: gap of unknown length
555304 556720: gap of 3605 bp in length
556721 558137: gap of unknown length
558138 559554: gap of 3618 bp in length
559555 560971: gap of unknown length
560972 562388: gap of 3631 bp in length
562389 563805: gap of unknown length
563806 565222: gap of 3644 bp in length
565223 566639: gap of unknown length
566640 568056: gap of 3657 bp in length
568057 569473: gap of unknown length
569474 570890: gap of 3670 bp in length
570891 572307: gap of unknown length
572308 573724: gap of 3683 bp in length
573725 575141: gap of unknown length
575142 576558: gap of 3696 bp in length
576559 577975: gap of unknown length
577976 579392: gap of 3709 bp in length
579393 580809: gap of unknown length
580810 582226: gap of 3722 bp in length
582227 583643: gap of unknown length
583644 585060: gap of 3735 bp in length
585061 586477: gap of unknown length
586478 587894: gap of 3748 bp in length
587895 589311: gap of unknown length
589312 590728: gap of 3761 bp in length
590729 592145: gap of unknown length
592146 593562: gap of 3774 bp in length
593563 594979: gap of unknown length
594980 596396: gap of 3787 bp in length
596397 597813: gap of unknown length
597814 599230: gap of 3800 bp in length
599231 600647: gap of unknown length
600648 602064: gap of 3813 bp in length
602065 603481: gap of unknown length
603482 604898: gap of 3826 bp in length
604899 606315: gap of unknown length
606316 607732: gap of 3839 bp in length
607733 609149: gap of unknown length
609150 610566: gap of 3852 bp in length
610567 611983: gap of unknown length
611984 613400: gap of 3865 bp in length
613401 614817: gap of unknown length
614818 616234: gap of 3878 bp in length
616235 617651: gap of unknown length
617652 619068: gap of 3891 bp in length
619069 620485: gap of unknown length
620486 621902: gap of 3904 bp in length
621903 623319: gap of unknown length
623320 624736: gap of 3917 bp in length
624737 626153: gap of unknown length
626154 627570: gap of 3930 bp in length
627571 628987: gap of unknown length
628988 630404: gap of 3943 bp in length
630405 631821: gap of unknown length
631822 633238: gap of 3956 bp in length
633239 634655: gap of unknown length
634656 636072: gap of 3969 bp in length
636073 637489: gap of unknown length
637490 638906: gap of 3982 bp in length
638907 640323: gap of unknown length
640324 641740: gap of 3995 bp in length
641741 643157: gap of unknown length
643158 644574: gap of 4008 bp in length
644575 645991: gap of unknown length
645992 647408: gap of 4021 bp in length
647409 648825: gap of unknown length
648826 650242: gap of 4034 bp in length
650243 651659: gap of unknown length
651660 653076: gap of 4047 bp in length
653077 654493: gap of unknown length
654494 655910: gap of 4060 bp in length
655911 657327: gap of unknown length
657328 658744: gap of 4073 bp in length
658745 660161: gap of unknown length
660162 661578: gap of 4086 bp in length
661579 662995: gap of unknown length
662996 664412: gap of 4099 bp in length
664413 665829: gap of unknown length
665830 667246: gap of 4112 bp in length
667247 668663: gap of unknown length
668664 670080: gap of 4125 bp in length
670081 671497: gap of unknown length
671498 672914: gap of 4138 bp
```

```

FEATURES
  source
    1..228527
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="RP11-115E23"
      /clone_1fb="PBC1-11"
BASE COUNT      69299 a 46162 c 45464 g 67423 t      179 others
ORIGIN
Query Match      43.4%  Score 137;  DB 22;  Length 228527;
Best Local Similarity 43.5%  Pctd. No. 1,43e+02;
Matches 30;  Conservative 14;  Mismatches 25;  Indels 0;  Gaps 0;

Db 60353 AGTAATTTTAAATTAAGGTGGATGCATGATGTCACATGTGTAACATGAAG 60412
OY 2  ARGARTTYTNACNCCNBARARVTCARTGYTGNGAYTTCAGTATWMSAAGAYG 61
Db 60413 TTTGAGCCC 60421
OY 62  TMTGYGCNC 70
          1 11111
          1 11111

RESULT      8
LOCUS      AC008969 135499 bp      DNA      HTG      18-FEB-2000
DEFINITION Homo sapiens chromosome 19 clone CTD-2368P22, WORKING DRAFT
ACCESSION   AC008969
VERSION     AC008969.3  GI:6997052
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominiidae; Homo.
            1 (bases 1 to 135499)
            DOE Joint Genome Institute.
            Sequencing of Human Chromosome 19
            Unpublished
            2 (bases 1 to 135499)
            DOE Joint Genome Institute.
            Direct Submission
            Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2600 Mitchell Drive, Walnut Creek, CA 94598, USA
            On Feb 18, 2000 this sequence version replaced gi:600941.
            -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----Summary Statistics
            Consensus quality: 98498 bases at least Q40
            Consensus quality: 108756 bases at least Q30
            Consensus quality: 112157 bases at least Q20
            Estimated insert size: 135499; sum-of-contigs estimation
            Estimated insert size: 98200; agarose-tip estimation
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 33 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            * 1 1019: contig of 1019 bp in length
            *
            * gap of unknown length

```

1020	2453:	contig of 134 bp in length
*		gap of unknown length
*	2454	contig of 1075 bp in length
*		gap of unknown length
*	3529	contig of 1349 bp in length
*		gap of unknown length
*	4878	contig of 1591 bp in length
*		gap of unknown length
*	6469	contig of 1048 bp in length
*		gap of unknown length
*	7517	contig of 1390 bp in length
*		gap of unknown length
*	8907	contig of 1103 bp in length
*		gap of unknown length
*	10010	contig of 1055 bp in length
*		gap of unknown length
*	11065	contig of 1053 bp in length
*		gap of unknown length
*	12118	contig of 1080 bp in length
*		gap of unknown length
*	13198	contig of 1015 bp in length
*		gap of unknown length
*	14213	contig of 1077 bp in length
*		gap of unknown length
*	15290	contig of 2052 bp in length
*		gap of unknown length
*	17342	contig of 1073 bp in length
*		gap of unknown length
*	18415	contig of 1956 bp in length
*		gap of unknown length
*	20371	contig of 1038 bp in length
*		gap of unknown length
*	21409	contig of 1086 bp in length
*		gap of unknown length
*	22495	contig of 1014 bp in length
*		gap of unknown length
*	22495	contig of 1014 bp in length
*		gap of unknown length
*	23509	contig of 2430 bp in length
*		gap of unknown length
*	25339	contig of 1011 bp in length
*		gap of unknown length
*	26950	contig of 1005 bp in length
*		gap of unknown length
*	27955	contig of 1583 bp in length
*		gap of unknown length
*	29538	contig of 1356 bp in length
*		gap of unknown length
*	30894	contig of 1249 bp in length
*		gap of unknown length
*	32143	contig of 1217 bp in length
*		gap of unknown length
*	33360	contig of 3483 bp in length
*		gap of unknown length
*	36843	contig of 6603 bp in length
*		gap of unknown length
*	43445	contig of 6537 bp in length
*		gap of unknown length
*	44982	contig of 6262 bp in length
*		gap of unknown length
*	49983	contig of 14541 bp in length
*		gap of unknown length
*	56245	contig of 26425 bp in length
*		gap of unknown length
*	70786	contig of 38285 bp in length
*		gap of unknown length
*	97211	contig of 38285 bp in length

FEATURES	SOURCE	Location/Qualifiers
		1: 135499
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="19"
		/clone="CTD-2368P22"
BASE COUNT	36143 a 33819 c 33220 g 33260 t	57 others
ORIGIN		
Query Match	42.7%	Score 135; DB 24; Length 135499;

[illegible]

RESULT	9				
LOCUS	AC014821	17438 bp	DNA		16-NOV-1999
DEFINITION	Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, In Ordered				
ACCESSION	AC014821				
VERSION	AC014821.1	GI:6436514			
KEYWORDS	HTG; HTGS_PHASE2.				
SOURCE	IruIt tly.				
ORGANISM	Drosophila melanogaster				

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (phases 1 to 17438)  
Adams, M. and Venter, J.C.  
Direct Submission  
Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA  
This sequence was identified as CDM:10210569 by the submitter.

COMMENTS  
This sequence was identified as CDM:U0210559 by the submitter.  
For further information on this sequence e-mail to [fly@celera.com](mailto:fly@celera.com).  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced.  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
Location/Qualifiers

```

/organism="Drosophila melanogaster
/db_xref="taxon:7227"
BASE COUNT      5022 a      3904 c      3790 g      4722 t
ORIGIN

```

	Query Match	41.8%	Score 132	Length 17438
	Best Local Similarity	45.1%	Pred. No. 2.55e+02	
	Matches	23	Mismatches 15	Indels 0
			Gaps	0
Db	16810	AAATCTAGTATCTGCACGTACGTGTATACCAACACCAAGTCCGCCCA	16860	
Oy	22	AATTACATGCTGNGAATTCAGATATATATTTTAAATGATATGTCGNCAR	72	

RESULT	10	AC012665	148162 bp	DNA	HTG	31-JAN-2000
LOCUS						
DEFINITION						
ACCESSION		AC012665				
VERSION		AC012665.3	GI:6838945			
KEYWORDS		HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE		human.				
ORGANISM		Homo sapiens				

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE
1	(bases 1 to 148162)	The sequence of Homo sapiens clone	unpublished	
2	(bases 1 to 148162)			

**AUTHORS** Waterston, R.H.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (02-NOV-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis MO 63108, USA  
**COMMENT** On Jan 31, 2000 this sequence version replaced g1:6358866.

```
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
```



[illegible]

	LOCUS	AC007777	151670 bp	DNA	HTG	10-JUN-1999
	DEFINITION	Homo sapiens chromosome 18 clone hRPK.493_O.9 map 18, ***				
	ACCESSION	AC007777.1	GI:5032340			
	VERSION	AC007777.1				
	KEYWORDS	HTG; HTGS_PHASE1.				
	SOURCE	human.				
	ORGANISM	Homo sapiens				
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
	AUTHORS	1 (bases 1 to 151670)				
	TITLE	Birren,B., Linton,L., Nusbaum,C. and Lander,E.				
	JOURNAL	Homo sapiens chromosome 18, clone hRPK.493_O.9 unpublished				
	REFERENCE	2 (bases 1 to 151670)				
	AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckely,R., Benn,J., Brown,A., Castle,A., Cenny,J., Colangelo,M., Collins,S., Collamore,A., Cooke,P., Dearlano,K., Depayre,E., Devon,K., Dewar,K., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Garduna,S., Gilbert,D., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L., Karakas,A., Lebecky,J., Liu,C., Locke,K., MacDonald,P., Margulis,N., McMan,P., McGuck,A., McKernan,K., McLaughlin,J., Meldrum,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Naylor,O., Nilot,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P., Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A., Taghvaei,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.				
	COMMENT	Direct Submission Submitted (10-JUN-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <a href="http://ftp.genome.washington.edu/RM/RepeatMasker.html">http://ftp.genome.washington.edu/RM/RepeatMasker.html</a> . * NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.				
	TITLE	1 29014: contig of 29014 bp in length				
	JOURNAL	gap of unknown length				
	COMMENT	29015 43358: contig of 14344 bp in length				
	TITLE	43359 62409: gap of unknown length				
	JOURNAL	contig of 19051 bp in length				
	COMMENT	62410 81825: gap of unknown length				
	TITLE	81826 139797: contig of 19416 bp in length				
	JOURNAL	gap of unknown length				
	COMMENT	139798 151670: contig of 57972 bp in length				
	TITLE	Location/Qualifiers				
	JOURNAL	1. 151670				
	FEATURES	/organism="Homo sapiens"				
	SOURCE	/db_xref="taxon:9606"				
		/clone="hRPK.493_O.9"				
		/clone_lib="RPCI-II human BAC library"				
		/map="18"				
		/chromosome="18"				
	BASE COUNT	48397 a 28102 c 27384 g 47255 t 532 others				
	ORIGIN					
	Query Match	41.8%; Score 132; DB 28; Length 151670;				
	Best Local Similarity	44.0%; Pred. No. 2,53e+02;				
	Matches	22; Conservative 13; Mismatches 15; Indels 0; Gaps 0;				
Db	59698	TTCTAATAGAGAATAATTATGACTGTAGCAITATTCGGAGACAAGAA 59747				
sp	56	TTTNSWDITNACRTGNARRCTNCACCATCATGAGARATTTTNGGNTNNRRA 7				

RESULT 13  
LOCUS AC008460 156492 bp DNA HTG 26-JAN-2000  
DEFINITION Homo sapiens chromosome 5 clone CTC-352D12, LOW-PASS SEQUENCE  
SAMPLING.  
ACCESSION AC008460  
VERSION AC008460.2 GI:6759081  
KEYWORDS HTG; HTGS\_PHASE0.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 156492)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 5  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 156492)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jan 26, 2000 this sequence version replaced gi:5686606.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
\* NOTE: This record contains 106 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
1  
\* 756: contig of 756 bp in length  
\* gap of unknown length  
\* 757 1086: contig of 330 bp in length  
\* gap of unknown length  
\* 1087 1739: contig of 653 bp in length  
\* gap of unknown length  
\* 1740 2862: contig of 1123 bp in length  
\* gap of unknown length  
\* 2863 3766: contig of 904 bp in length  
\* gap of unknown length  
\* 3767 4535: contig of 769 bp in length  
\* gap of unknown length  
\* 4536 5747: contig of 1212 bp in length  
\* gap of unknown length  
\* 5748 5900: contig of 153 bp in length  
\* gap of unknown length  
\* 5901 6346: contig of 446 bp in length  
\* gap of unknown length  
\* 6347 7132: contig of 786 bp in length  
\* gap of unknown length  
\* 7133 7576: contig of 444 bp in length  
\* gap of unknown length  
\* 7577 7799: contig of 223 bp in length  
\* gap of unknown length  
\* 7800 8393: contig of 594 bp in length  
\* gap of unknown length  
\* 8394 9096: contig of 703 bp in length  
\* gap of unknown length  
\* 9097 9994: contig of 898 bp in length  
\* gap of unknown length  
\* 9995 10596: contig of 602 bp in length  
\* gap of unknown length  
\* 10597 11598: contig of 1002 bp in length  
\* gap of unknown length


11599 12565: contig of 967 bp in length  
\* gap of unknown length  
\* 12566 13291: contig of 726 bp in length  
\* gap of unknown length  
\* 13292 14019: contig of 728 bp in length  
\* gap of unknown length  
\* 14020 14942: contig of 923 bp in length  
\* gap of unknown length  
\* 14943 15638: contig of 696 bp in length  
\* gap of unknown length  
\* 15639 16562: contig of 924 bp in length  
\* gap of unknown length  
\* 16563 17495: contig of 933 bp in length  
\* gap of unknown length  
\* 17496 18590: contig of 1095 bp in length  
\* gap of unknown length  
\* 18591 19370: contig of 780 bp in length  
\* gap of unknown length  
\* 19371 21061: contig of 1691 bp in length  
\* gap of unknown length  
\* 21062 21661: contig of 600 bp in length  
\* gap of unknown length  
\* 21662 22338: contig of 677 bp in length  
\* gap of unknown length  
\* 22339 23280: contig of 942 bp in length  
\* gap of unknown length  
\* 23281 24811: contig of 1531 bp in length  
\* gap of unknown length  
\* 24812 25884: contig of 1073 bp in length  
\* gap of unknown length  
\* 25885 26707: contig of 823 bp in length  
\* gap of unknown length  
\* 26708 27789: contig of 1082 bp in length  
\* gap of unknown length  
\* 27790 29117: contig of 1328 bp in length  
\* gap of unknown length  
\* 29118 30118: contig of 1001 bp in length  
\* gap of unknown length  
\* 30119 30860: contig of 742 bp in length  
\* gap of unknown length  
\* 30861 32125: contig of 1265 bp in length  
\* gap of unknown length  
\* 32126 33293: contig of 1168 bp in length  
\* gap of unknown length  
\* 33294 34626: contig of 1333 bp in length  
\* gap of unknown length  
\* 34627 35841: contig of 1215 bp in length  
\* gap of unknown length  
\* 35842 36803: contig of 962 bp in length  
\* gap of unknown length  
\* 36804 37840: contig of 1037 bp in length  
\* gap of unknown length  
\* 37841 38839: contig of 999 bp in length  
\* gap of unknown length  
\* 38840 39939: contig of 1100 bp in length  
\* gap of unknown length  
\* 39940 40973: contig of 1034 bp in length  
\* gap of unknown length  
\* 40974 41982: contig of 1009 bp in length  
\* gap of unknown length  
\* 41983 42857: contig of 875 bp in length  
\* gap of unknown length  
\* 42858 44198: contig of 1341 bp in length  
\* gap of unknown length  
\* 44199 45225: contig of 1027 bp in length  
\* gap of unknown length  
\* 45226 46646: contig of 1421 bp in length  
\* gap of unknown length  
\* 46647 48279: contig of 1633 bp in length  
\* gap of unknown length  
\* 48280 49444: contig of 1165 bp in length  
\* gap of unknown length  
\* 49445 51012: contig of 1568 bp in length













(TM)

misc_feature	93..113	/note="propeptide"
mat_peptide	114..824	/product="prostate specific antigen"
BASE COUNT	209 a 291 c 273 g 217 t	
ORIGIN		
Query Match	98.7%; Score 312; DB 1; Length 990;	
Best Local Similarity	60.8%; Pred. No. 2,23e-16;	
Matches	45; Conservative 19; Mismatches 10; Indels 0; Gaps 0	
Db	528 GAGGAGTCTTGACCCCAAGAAACTTCAGTGTGTGACCTCCATGTTATTTCCATGAC	587
Oy	1 GARGARTTYTNACNCCNARARATNTCARFTGYGTNGAYTYNCAVGTNATHTMSNAAYGAY	60
Db	588 GTGTGTGCCCAAGT 601	
Oy	61 GTNTGYCGNCARGT 74	
RESULT	2	
LOCUS	167863	992 bp DNA PAT 04-FEB-1998
DEFINITION	Sequence 13 from patent US 5674682.	
ACCESSION	167863	
VERSION	I67863.1 GI:282985	
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 992)	
AUTHORS	Croce,C., Gomella,L., Mulholland,S.Grnt, Moreno,J.G. and Fischer,R.	
TITLE	Nucleic acid primers for detecting micrometastasis of prostate cancer	
JOURNAL	Patent: US 5674682-A 13 07-OCT-1997;	
FEATURES	Location/Qualifiers	
source	1..992	
BASE COUNT	210 a 291 c 272 g 217 t 2 others	
ORIGIN		
Query Match	98.7%; Score 312; DB 1; Length 992;	
Best Local Similarity	60.8%; Pred. No. 2,23e-16;	
Matches	45; Conservative 19; Mismatches 10; Indels 0; Gaps 0	
Db	530 GAGGAGTCTTGACCCCAAGAAACTTCAGTGTGTGACCTCCATGTTATTTCCATGAC	589
Oy	1 GARGARTTYTNACNCCNARARATNTCARFTGYGTNGAYTYNCAVGTNATHTMSNAAYGAY	60
Db	590 GTGTGTGCCCAAGT 603	
Oy	61 GTNTGYCGNCARGT 74	
RESULT	3	
LOCUS	HUMPA 1415 bp mRNA PRI 07-JAN-1995	
DEFINITION	Human prostate specific antigen (PA) gene, 3' end, clone PA 75.	
ACCESSION	M21895	
VERSION	M21895.1 GI:189523	
KEYWORDS	glycoprotein; prostate antigen; prostate-specific antigen; seminogelin.	
SOURCE	Human prostate tumor cell line PC 82, cDNA to mRNA, clone PA 75.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
AUTHORS	Riegman,P.H., Klaassen,P., van der Korput,J.A., Romijn,J.C. and Trapman,J.	
TITLE	Molecular cloning and characterization of novel prostate antigen cDNA's	
JOURNAL	Biochem. Biophys. Res. Commun. 155 (1), 181-188 (1988)	
FEATURES	Location/Qualifiers	

source	1. .1415	/organism="Homo sapiens"
	/db_xref="taxon:9606"	
	/map="19q13.3"	
sig_peptide	<1. .40	/gene="Aps"
	/note="prostate specific antigen signal peptide"	
CDS	<1. .776	/gene="Aps"
	/note="prostate specific antigen precursor"	
	/codon_start=2	
	/db_xref="GDB:G00-119-695"	
	/protein_id="AA59995.1"	
	/db_xref="GI:189524"	
	/translation="VFPLTISVWIGAPILISRIVGMECEKHSOPWQVLVASRRIA KQGVLVHPQWLVLAHCIRKNSVILIGRSLPHPEPTGVFQVSHSFPLDMSLL KNRLRPGDSSHDLMILRLSEPAELIDAKVMDLPTQEPALGTCYASGWSIEPEE FLIPRKLCQCDVLHVISNDVCAQVHPQVTKFMDCAGMTGKSGSGSGCPVNCV LQGTWSSEPCALPERPSLYTKVHVHRKMKIDTVANP"	
gene	1. .775	/gene="Aps"
	62. .772	/gene="Aps"
mat_peptide	/note="prostate specific antigen"	
BASE COUNT	328 a 361 c 410 g 316 t	
ORIGIN	176 bp upstream of PvuII site.	
Query Match	98.7%; Score 312; DB 6; Length 1415;	
Best Local Similarity	60.8%; Pred. No. 2,236-16;	
Matches	45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;	
Db	476 GAGGAGTCTTGACCCCAAGAACTCAGTGTGGACCTCCATGTATTCCATGAC 535	
QY	1 GARGARTYYTNACCCAAARAARAYTNCAARGTGGCAYGYNATWSMAAYGAY 60	
Db	536 GTGTGTGGCAAGT 549	
QY	61 GTNTGCGNCARGT 74	
RESULT	4	
LOCUS	HUMAPS 1446 bp mRNA PRI 04-JAN-1995	
DEFINITION	Homo sapiens prostate-specific antigen mRNA, complete cds.	
ACCESSION	M26663	
VERSION	M26663.1 GI:618463	
KEYWORDS	prostate-specific antigen.	
SOURCE	Homo sapiens prostate gland CDNA to mRNA.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
	Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE	1 (sites)	
AUTHORS	Henttu,P. and Viikho,P.	
TITLE	cDNA coding for the entire human prostate specific antigen shows	
JOURNAL	high homologies to the human tissue kallikrein genes	
MEDLINE	Biochem. Biophys. Res. Commun. 160 (2), 903-910 (1989)	
COMMENT	8924651	
	On Jan 5, 1995 this sequence version replaced gi:341511.	
FEATURES	Ref [1] reports bases 1. .135 only.	
source	Location/Qualifiers	
	1. .1446	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="PSA-20"	
	/tissue_type="prostate gland"	
	/map="19q13.3"	
	1. .24	
	/gene="Aps"	
	/note="G00-119-695"	
gene	1. .1446	
	/gene="Aps"	
CDS	25. .810	
	/gene="Aps"	
	/note="AA 18. .24 propeptide; AA 69 glycosylated Asn	

```

residue
/codon_start=1
/db_xref="GDB:G00-119-695"
/product="prostate-specific antigen"
/protein_id="AA38802.1"
/db_xref="GI:618464"
/translation="MWPFVFLTSLVWIGAPLILSRIVGMECEKHSOPMOVLVAS
NGRAVCGGVLPQWVLTAAHCIRKNSVILGRSLFPEPTGQVFSHSPFLYD
MSLKNRFLRPEDSSHDMLRLSEPAELDAYVMDLPTQEPALGTTCYASGSGSI
EPEFLTRPKLCQVDLHVISNDVCAQVHPQVTKFMLCAGKMTGKSTCSGSGPLV
CNGVLGITSMSSEPCALPERSLTKVHYHRKMKITIVANP"
sig_peptide
25..75
/gene="APS"
/organism="Homo sapiens"
/note="G00-119-695"
mat_peptide
94..807
/gene="APS"
/note="G00-119-695"
/product="prostate-specific antigen"
3'UTR
811..1446
/gene="APS"
/note="G00-119-695"
polyA_site
1446
/gene="APS"
/note="G00-119-695"
BASE COUNT 333 a 372 c 419 g 322 t
ORIGIN
Query Match 98.7%; Score 312; DB 6; Length 1446;
Best Local Similarity 60.8%; Pred. No. 2,236-16;
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;
Db 511 GAGGAGTCTTGACCCCAAGAAGAACTTCAGTGTGACCTCATGTTATTTCCAAATGAC 570
Oy 1 GARGARTTYTNACNCNNAARAARTNCARTGTGNGAYITNCAGTATNTHMSNAAYGAY 60
Db 571 GGTGTGCGCAAGT 584
Oy 61 GTNTGYCNCARGT 74
LOCUS 5 167864 1462 bp DNA PAT 04-FEB-1998
DEFINITION Sequence 14 from patent US 5674682.
ACCESSION 167864
VERSION 167864.1 GI:2829986
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1462)
Croce, C., Gomella, L., Mulholland, S., Grant, Moreno, J. G. and
Fischer, R.
Nucleic acid primers for detecting micrometastasis of prostate
cancer
Patent: US 5674682-A 14 OCT-1997;
LOCATION/Qualifiers
1..1462
BASE COUNT 343 a 374 c 422 g 323 t
ORIGIN
Query Match 98.7%; Score 312; DB 1; Length 1462;
Best Local Similarity 60.8%; Pred. No. 2,236-16;
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;
Db 511 GAGGAGTCTTGACCCCAAGAAGAACTTCAGTGTGACCTCATGTTATTTCCAAATGAC 570
Oy 1 GARGARTTYTNACNCNNAARAARTNCARTGTGNGAYITNCAGTATNTHMSNAAYGAY 60
Db 571 GGTGTGCGCAAGT 584
Oy 61 GTNTGYCNCARGT 74

```

```

RESULT 6
LOCUS HSPSAR 1466 bp mRNA PRI 21-MAR-1995
DEFINITION Human mRNA for prostate specific antigen.
ACCESSION X05332
VERSION X05332.1 GI:35740
KEYWORDS
antigen; kallikrein-like protein; prostate specific antigen; signal
peptide.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 1415)
Lundvall, A. and Lilja, H.
Molecular cloning of human prostate specific antigen cDNA
FEBS Lett. 214 (2), 317-322 (1987)
87190978
REFERENCE
2 (bases 1 to 1466)
Lundvall, A.
Direct Submission
Submitted (07-NOV-1989) to the EMBL/Genbank/DBJ databases
Data kindly reviewed (01-DEC-1987) by Lundvall A.
LOCATION/Qualifiers
1..1466
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="prostate"
/clone_lib="(lambda)gt11"
/clone="lambda HPSA-1"
1..1466
/note="primary transcript"
44..829
/note="precursor polypeptide (AA -17 to 244)"
/codon_start=1
/protein_id="CAA28947.1"
/db_xref="GI:35741"
/db_xref="SWISS-PROT:P07288"
/translation="MWPFVFLTSLVWIGAPLILSRIVGMECEKHSOPMOVLVAS
NGRAVCGGVLPQWVLTAAHCIRKNSVILGRSLFPEPTGQVFSHSPFLYD
MSLKNRFLRPEDSSHDMLRLSEPAELDAYVMDLPTQEPALGTTCYASGSGSI
EPEFLTRPKLCQVDLHVISNDVCAQVHPQVTKFMLCAGKMTGKSTCSGSGPLV
CNGVLGITSMSSEPCALPERSLTKVHYHRKMKITIVANP"
44..94
/product="signal peptide (AA -17 to -1)"
95..115
/note="propeptide (AA 1-7)"
116..826
/product="mature prostate specific antigen"
1445..1450
/note="put. polyA signal"
1466
/note="polyA site"
382 c 422 g 324 t
BASE COUNT 338 a 382 c 422 g 324 t
ORIGIN
Query Match 98.7%; Score 312; DB 7; Length 1466;
Best Local Similarity 60.8%; Pred. No. 2,236-16;
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;
Db 530 GAGGAGTCTTGACCCCAAGAAGAACTTCAGTGTGACCTCATGTTATTTCCAAATGAC 589
Oy 1 GARGARTTYTNACNCNNAARAARTNCARTGTGNGAYITNCAGTATNTHMSNAAYGAY 60
Db 590 GGTGTGCGCAAGT 603
Oy 61 GTNTGYCNCARGT 74
RESULT 7
LOCUS HUMPA8 1654 bp mRNA PRI 07-JAN-1995
DEFINITION Human prostate specific antigen (PA) gene, 3' end, clone PA 525.
ACCESSION M21896
VERSION M21896.1 GI:189525

```



VERSION	X73560.1	GI:311843
KEYWORDS	chymotrypsin-like; glycosylation; kallikrein-like protein; prostatic specific antigen; prostatic secretory protein.	
SOURCE	rhesus monkey.	
ORGANISM	Macaca mulatta	
REFERENCE	Eukaryotes: Metazoa: Chordata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Cercopithecoidea: Cercopithecinae; Macaca.	
AUTHORS	Gauthier,E.R., Chapdelaine,P., Tremblay,R.R. and Dube,J.Y.	
TITLE	Characterization of rhesus monkey prostate specific antigen cDNA	
JOURNAL	Biochim. Biophys. Acta 1174 (2), 207-210 (1993)	
MEDLINE	93363642	
REFERENCE	2 (bases 1 to 1514)	
AUTHORS	Dube,J.Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (20-Apr-1993) J.Y. Dube, Laval University Hospital Centre, Lab. of Hormonal Bioregulation, 2705 Laurier Boulevard, Ste-Foy, Quebec, G1V 4G2, CANADA	
FEATURES	Location/Qualifiers	
source	1..1514	
CDS	/organism="Macaca mulatta"	
	/db_xref="taxon:9544"	
	/dev_stage="adult"	
	/clone_lib="Rhesus monkey prostate cDNA"	
	33..818	
	/codon_start=1	
	/product="prostate specific antigen"	
	/protein_id="CAA51937.1"	
	/db_xref="GI:311844"	
	/db_xref="SWISS-PROT:P33619"	
	/translation="MWLVVFETLSVTNLTGAPLIRSYIVSGMECEKHSOPMOVLVA KRAVGCVLHPQWYLVLAHCIRENSVILIGRNHPPYPEDMGQFYSHSPHLXYA MSLNRLYLPGDDSDHDMLRLERLEPEITDAVOLPLTERELGTICRASGGSGS EPEHLTRFKLCQCDLHIISDVCAGVSHQRKYTKMLCAGSMWGSKSTCSGSBSGPL CDGVQGITSWSQCPCALPRRSPLYTKVVRKKMIODTIMANP"	
s1g_peptide	33..107	
mat_peptide	105..815	
misc_feature	336..338	
	/note="N-glycosylation site"	
BASE COUNT	341 a 393 c 445 g 335 t	
ORIGIN		
Query Match	89.2% Score No.2 DB 7 Length 1514;	
Best Local Similarity	56.8%; Pred.No.2,18-13:	
Matches	42; Conservative 19; Mismatches 13; Indels 0; Gaps 0;	
Db	519 GAGGAACACTTGACTCCAAAGAAGCTCACTGTGTGGACCTCATATTTCCAAATGAT 578   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   QY 1 GAGGARTTYTNACCNNARAARTNCARTGYGTNGAYTCAYTNCAYGNATHMSMAAYGAY 60	
Db	579 GTGTGTGCCGCAAT 592    :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
QY	61 GTNTGYGCNCARGT 74	
LOCUS	11 HSPSA4 281 bp DNA PRI 24-FEB-1999	
DEFINITION	Human psa gene for prostate specific antigen exon 4.	
ACCESSION	X13943	
VERSION	X13943.1 GI:35728	
KEYWORDS	kallikrein; prostate specific antigen; psa gene.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia; Eutheria: Primates: Catarrhini: Homnidae; Homo.	
REFERENCE	1 (bases 1 to 281)	
AUTHORS	Digby,M.R.	
TITLE	Direct Submission	
JOURNAL	Submitted (09-JAN-1989) Digby M.R., The Howard Florey Institute, Melbourne University, Parkville, Victoria 3052, Australia	
AUTHORS	2 (bases 1 to 281) Digby,M., Zhang,X.Y. and Richards,R.I.	
TITLE	Human prostate specific antigen (PSA) gene: structure and linkage	

```

JOURNAL      to the kallikrein-like gene, hK2-1
MEDLINE      89183632
COMMENT      Nucleic Acids Res. 17 (5), 2137 (1989)
FEATURES
  source      Data kindly reviewed (16-may-1989) by Digby M.R.
              Location/Qualifiers
                1..281
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="19"
                  /clone_1db="cosmid"
                  /clone="CHK-2"
                  <1..52
                    /number=3
                    53..189
                      /number=4
                      <53..>189
                        /note="prostate specific antigen"
                        /codon_start=3
                        /protein_id="CAA32126.1"
                        /db_xref="GI:1335283"
                        /db_xref="SWISS-PROT:P07288"
                        /translation="LTPKKLQCDLHVISNDVCAQVHPQKVTKEFMLCAGRTGKSTC
                          S"
                          190..>281
                            /number=4
                            64 a      71 c      82 g      64 t
BASE COUNT
ORIGIN
Query Match      87.7%; Score 277; DB 7; Length 281;
Best Local Similarity 59.7%; Pred. No. 6,78e-13;
Matches 40; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

Db 53 TCTTGACCCCAAGAAATTCAGTGTGGACCTCCAGTATTTCCATGAGCTGTGG 112
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 8 TTTTACNCCNAARARYTCARTGTGTGATTCATGATNATHWSNAATGATGTGIG 67

Db 113 CGCAGT 119
      ||:||||
QY 68 CMCAGT 74

RESULT 12
LOCUS      HSPSAG 5873 bp DNA PRI 24-APR-1993
DEFINITION Human DNA for prostate specific antigen (PSA).
ACCESSION  X14810
VERSION  X14810.1 GI:35732
KEYWORDS   Ali repetitive sequence; prostate specific antigen; serine
            protease.
SOURCE      human.
            ORGANISM Homo sapiens
                Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
                Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 5873)
            Klobbeck,H.G., Combratito,G., Schulz,P., Arbusow,V. and Fittler,F.
            Genomic sequence of human prostate specific antigen (PSA)
            Nucleic Acids Res. 17 (10), 3981 (1989)
            89282407
            See also acc# X05332 and x07730.
            Data kindly reviewed (19th June 1989) by Klobbeck H.-G.
FEATURES
  source      Location/Qualifiers
                1..5873
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /cell_type="Lymphoid cell line GM 607"
                  332..338
                    /cell_type="Lymphoid cell line GM 607"
                    355..365
                      /note="transcriptional start region"
                      join(401..446,1688..1847,3477..3763,3907..4043,5413..5568)
                  /product="prostate specific antigen (PSA)"
                  /protein_id="CAA32915.1"
                  /db_xref="GI:296671"
                  /db_xref="SWISS-PROT:P07288"

```

	exon		<401. .446 /number=1 join<401. .446,1688. .1713> .5413. >.5568) join(401. .446,1688. .1713)
	mRNA		
	sig_peptide		
	intron		/number=1 1688. .1847 /number=2 join(1714. .1847,3477. .3763,3907. .4043,5413. .5565) /product="prostate specific antigen (PSA)"
	mat_peptide		1848. 3476 /number=2 2311. .2663 /note="Alu repetitive sequence"
	misc_feature		3477. .3763 /number=3 3764. .3906 /number=3 3907. .4043 /number=4 4044. .5412 /number=4 5413. .>5568 /number=5
	exon		
	intron		
	exon		
BASE COUNT	1186 a	1778 c	1503 g
ORIGIN	1778 c	1503 g	1406 t
Query Match	87.7%;	Score 277;	DB 7;
Best Local Similarity	59.7%;	Pred. No. 6.78e-13;	Length 5873;
Matches	40; Conservative	17; Mismatches	10; Indels 0; Gaps 0;
Db	3907	TCTTACCCCAAGAAGAACTGATGTGGTGAGCCTCGACGTATTTCACAATACGTTGTG	3966
Oy	8	TYTNACMCNARARYNTCARITYGTNGAYTNCAGYATNAHWSNAGAYGINTGYG	67
Db	3967	CGCAGT	3973
Oy	68	CNCARGT	74
RESULT	13		
LOCUS	HUMPSANTIG	6153 bp	DNA
DEFINITION	Human prostate-specific antigen (Pa) gene, complete cds.		PRI 08-JAN-1995
ACCESSION	M24543		
VERSION	M24543.1	GI:341200	
KEYWORDS	kallikrein; prostate-specific antigen.		
SOURCE	Homo sapiens	(tissue library: of G.Grosveld)	DNA.
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS	Eutheria; Primates; Catarrhini; Homidae; Homo.		
TITLE	Riegman,P.H., Vliestra,R.J., van der Korp,J.A., Romijn,J.C. and		
JOURNAL	Characterization of the prostate-specific antigen gene: a novel		
MEDLINE	human kallikrein-like gene		
FEATURES	Biochem. Biophys. Res. Commun. 159 (1), 95-102 (1989)		
SOURCE	89165891		
	Location/Qualifiers		
	1..6153		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/tissue_lib="of G.Grosveld"		
	71..84		
	/gene="PA"		
	/note="transcriptional regulatory element"		
	188..193		
	/gene="PA"		

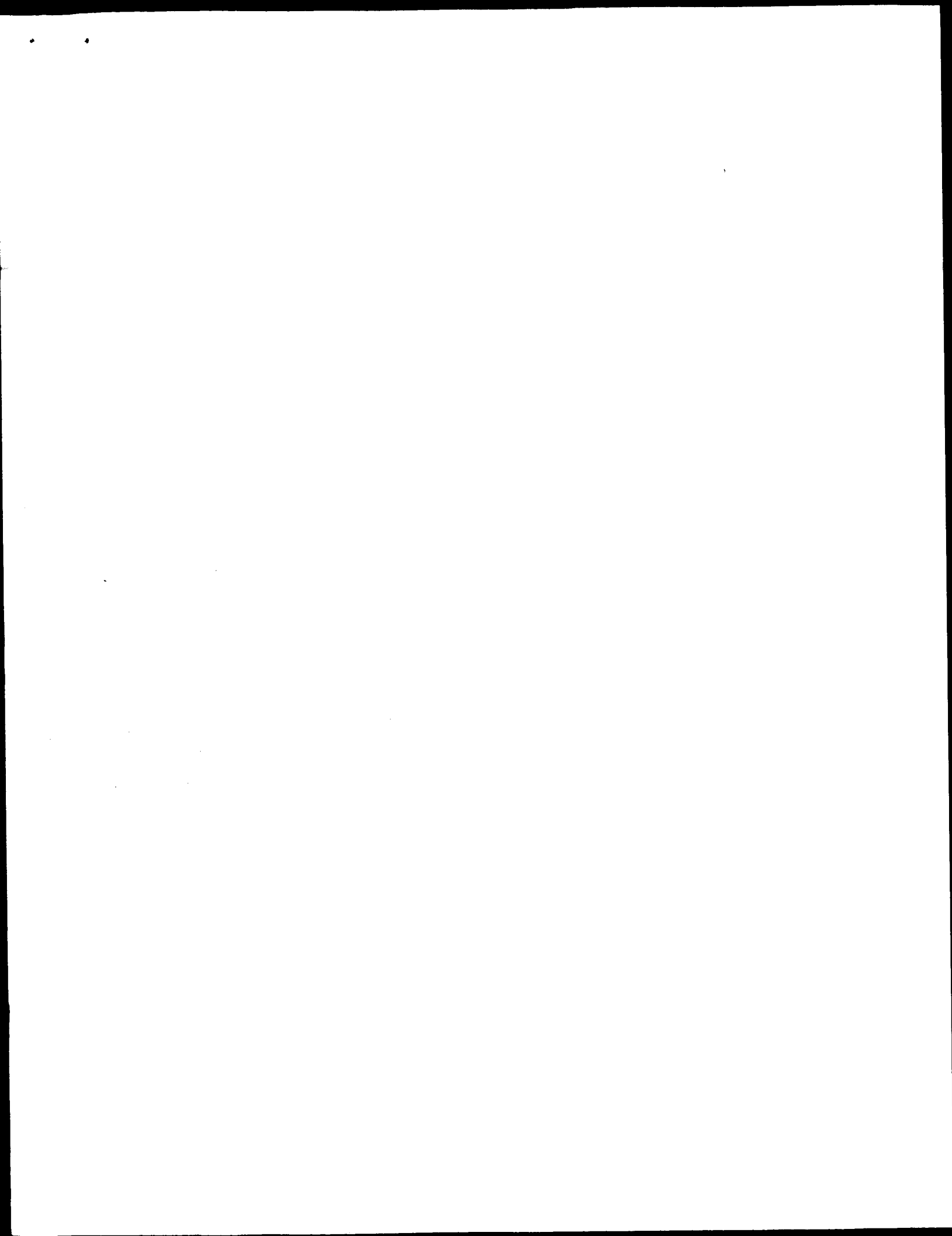
LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
LOCUS	HUMPSA	7130 bp	DNA	PRI	08-JAN-1995
DEFINITION	Human prostate specific antigen gene, complete cds.				
ACCESSION	M21274				
VERSION	M21274.1	GI:190552			
KEYWORDS	Alu repeat; kallikrein; prostate-specific antigen.				
SOURCE	Human peripheral leukocyte DNA, (library of R.Wydro), clones				
	lambda-PSA.[24A,25A,25B].				



[illegible]

	RESULT	15	A37262	7130 bp	DNA	PAT	05-MAR-1997
	LOCUS						
	DEFINITION		Sequence 2 from Patent WO9404196.				
	ACCESSION		A37262				
	VERSION		A37262.1	GI:2294359			
	KEYWORDS		.				
	SOURCE		. unidentified.				
	ORGANISM		. unidentified				
	REFERENCE		unclassified.				
	AUTHORS		1 (bases 1 to 7130)				
	TITLE		Ville,R.G. and Hart,I.R.				
	JOURNAL		TUMOUR THERAPY				
			Patent: WO 9404196-A 2 03-MAR-1994;				
			IMP CANCER RES TECH (GB)				
	FEATURES		Location/Organism				
	source		l.. 7130				
			/organism="unidentified"				
			/db_xref="taxon:32644"				
BASE COUNT		1530 a	2024 c	1867 g	1709 t		
ORIGIN							
Query Match		Best Local Similarity	87.7%:	Score 277,	DB 1; Length 7130;		
Matches	40;	Conservative	17;	Mismatches 10;	Indels 0; Gaps 0;		
Db	4186	TCCTGACCAGCAAAACAATTCCAGTGTGGTGACGCCTCATTTATTGCCAATGACTGTGCGT	G	4245			
Oy	8	TYTNACNCACAAAANAAYTNCARTGYCTNGAYTNCAYTNCATGNATHWSNMAAYGAATGTNYG	C	67			
Dd	4246	CGCAACT	4252				
Oy	68	CNCAGT	74				

Search completed: Fri Jun 23 17:20:36 2000  
Job time : 184 secs.







DR WP: 98-263285/24.  
 PT New transfected prostate cancer cells harbouring recombinant DNA  
 PT expression vector and expressing prostate-specific antigen - used to  
 PT treat and diagnose prostate cancer  
 PS Claim 5: Page 6-7: 19pp: English.  
 CC This polynucleotide comprises a DNA sequence coding for  
 CC propeptide-deleted human prostate specific antigen (PSA). It was  
 CC obtained by PCR amplification (see V32500-01). The PSA DNA was  
 CC cloned into expression vector pRC/CMV to yield vector  
 CC pRC/CMV-PSA-EPRO. A full-length wild-type PSA construct (see  
 CC W32496), pRC-CMV-PSA-WT, was also produced. Claimed transfectant  
 CC prostate cancer host cells harbour one of the novel recombinant  
 CC pRC-CMV vectors. Also claimed is evaluation of the effectiveness  
 CC of a test compound for the treatment or prevention of a condition  
 CC associated with an increase in PSA, comprising transfecting a  
 CC prostate cancer cell with an expression vector containing DNA  
 CC encoding PSA or modified PSA, culturing the cell under conditions  
 CC such that the PSA or modified PSA protein is expressed, exposing  
 CC the cell to a test compound and measuring the change in PSA  
 CC proteolytic activity relative to a control in which the transfected  
 CC cell is not exposed to a test compound. The cells and vectors are  
 CC used for the diagnosis, prognosis, treatment and study of prostate  
 CC cancer.  
 SQ Sequence 1445 BP; 336 A; 372 C; 418 G; 319 T;  
 Query Match 98.7%; Score 312; DB 1; Length 1445;  
 Best Local Similarity 60.8%; Pred. No. 2,22e-15;  
 Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;  
 Db 509 GAGGAGTTCTTGTACCCCAAGAACTTGTGAGCCCTCATGTTATTTCCATGAC 568  
 QY 1 GARGARTTYTTCACCCNAAARATTCARTGTGTGATTCATGTTATTCATGAC 60  
 Db 569 GTGTGTGCGCAAGT 582  
 QY 61 GTTGTGCGCAAGT 74  
 RESULT 4  
 ID T91055 standard; DNA: 1462 BP.  
 AC T91055.  
 DT 26-FEB-1998 (first entry)  
 DE Human prostate specific antigen gene probe PSA-20.  
 KW Prostate specific antigen; primer; human; prostate cancer;  
 KW metastasis; diagnosis; probe; ss.  
 OS Homo sapiens.  
 PN US5674682-A.  
 PD 07-OCT-1997.  
 PF 29-OCT-1992; 973322.  
 PR 15-DEC-1994; US-358782.  
 PR 29-OCT-1992; US-973322.  
 PR 23-AUG-1994; US-294611.  
 PA (UJYE-) UNIV JEFFERSON THOMAS.  
 PI Croce C, Fischer R, Gonnella L, Moreno JG, Mulholland SG;  
 DR WP: 97-502328/46.  
 PT Oligonucleotide primers specific for prostate specific antigen -  
 PT used for the diagnosis of prostate cancer micro-metastasis  
 PS Disclosure: Fig 5A-5B; 18pp: English.  
 CC This DNA sequence comprises PSA-20, a probe used in the detection  
 CC of amplified products of the human prostate specific antigen (PSA)  
 CC gene (see T91054). A method of detecting prostate cancer  
 CC micrometastasis involves: obtaining a sample of nucleic acids from  
 CC a patient tissue sample; amplifying nucleic acids specific for  
 CC prostate cancer (see also T91044-53) or amplifying a signal  
 CC generated by hybridisation of a probe specific to a prostate  
 CC cancer-specific nucleic acid in the sample; and detecting the  
 CC presence of amplified nucleic acids or amplified signal, this being  
 CC indicative of micrometastasis. The invention also includes a  
 CC method of detecting cells which express prostate cancer specific  
 CC sequences by detecting the presence of nucleic acids specific for  
 CC prostate cancer or a signal specific for prostate cancer.  
 SQ Sequence 1462 BP; 343 A; 374 C; 422 G; 323 T;

Query Match 98.7%; Score 312; DB 1; Length 1462;  
 Best Local Similarity 60.8%; Pred. No. 2,22e-15;  
 Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;  
 Db 511 GAGGAGTTCTTGTACCCCAAGAACTTGTGAGCCCTCATGTTATTTCCATGAC 570  
 QY 1 GARGARTTYTTCACCCNAAARATTCARTGTGTGATTCATGTTATTCATGAC 60  
 Db 571 GTGTGTGCGCAAGT 584  
 QY 61 GTTGTGCGCAAGT 74  
 RESULT 5  
 ID V32496 standard; DNA: 1466 BP.  
 AC V32496;  
 DT 28-SEP-1998 (first entry)  
 DE Prostate specific antigen wild-type DNA.  
 KW Prostate specific antigen; human; vector; pRC/CMV-PSA-WT;  
 KW prostate cancer; diagnosis; therapy; ss.  
 OS Homo sapiens.  
 FH Key  
 FT Location/Qualifiers  
 FT misc\_feature 13..826  
 FT /tag= "a  
 FT /note= "nucleotides 13-826 are preferred for  
 FT expression vector construction"  
 EP 843006-A2.  
 PD 20-MAY-1998.  
 PF 12-NOV-1997; 309108.  
 PR 15-NOV-1996; US-030947.  
 PA (ELIT.) LITLY & CO. ELI.  
 PI Dixon EP, Little SP, Sutkowski DM;  
 DR WP: 98-263285/24.  
 PT New transfected prostate cancer cells harbouring recombinant DNA  
 PT expression vector and expressing prostate-specific antigen - used to  
 PT treat and diagnose prostate cancer  
 PS Claim 1: Page 5-6: 19pp: English.  
 CC This polynucleotide comprises a wild-type DNA sequence coding for  
 CC human prostate specific antigen (PSA), including the signal peptide  
 CC and propeptide region. It was obtained by PCR amplification of  
 CC human prostate cDNA using primers (see V32498-99) based on a  
 CC reported human PSA cDNA sequence. The PSA DNA was cloned into  
 CC expression vector pRC/CMV to yield vector pRC/CMV-PSA-WT. A  
 CC propeptide-deleted construct (see W32497), pRC-CMV-PSA-EPRO, was  
 CC also produced. Claimed transfectant prostate cancer host cells  
 CC harbour one of the novel recombinant pRC-CMV vectors. Also claimed  
 CC is evaluation of the effectiveness of a test compound for the  
 CC treatment or prevention of a condition associated with an increase  
 CC in PSA, comprising transfecting a prostate cancer cell with an  
 CC expression vector containing DNA encoding PSA or modified PSA,  
 CC culturing the cell under conditions such that the PSA or modified  
 CC PSA protein is expressed, exposing the cell to a test compound and  
 CC measuring the change in PSA proteolytic activity relative to a  
 CC control in which the transfected cell is not exposed to a test  
 CC compound. The cells and vectors are used for the diagnosis,  
 CC prognosis, treatment and study of prostate cancer.  
 SQ Sequence 1466 BP; 338 A; 382 C; 422 G; 324 T;  
 Query Match 98.7%; Score 312; DB 1; Length 1466;  
 Best Local Similarity 60.8%; Pred. No. 2,22e-15;  
 Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;  
 Db 530 GAGGAGTTCTTGTACCCCAAGAACTTGTGAGCCCTCATGTTATTTCCATGAC 589  
 QY 1 GARGARTTYTTCACCCNAAARATTCARTGTGTGATTCATGTTATTCATGAC 60  
 Db 590 GTGTGTGCGCAAGT 603  
 QY 61 GTTGTGCGCAAGT 74  
 RESULT 6  
 ID T35867 standard; cDNA: 1728 BP.

AC T35867; (first entry)  
 DT 21-OCT-1996 (first entry)  
 DE Prostate-specific antigen (PSA) cDNA.  
 KW Prostate-specific antigen; metastasis; carcinoma; tumour; cancer;  
 KM neoplasm; archival tissue; polymerase chain reaction; RT-PCR;  
 OS Homo sapiens.  
 FH Key  
 FT cds Location/Qualifiers  
 FT 68..82  
 FT /\*tag= a  
 FT 354..1092  
 FT /\*tag= b  
 FT mat\_peptide 378..1089  
 FT /\*tag= c  
 FT misc\_difference 752  
 FT /\*tag= d  
 FT /note= "an additional g residue is inserted  
 FT after position 752 in the sequence  
 FT given on page 39 of the specification"  
 FT 1706..1711  
 FT /\*tag= e  
 FT poly\_a\_signal  
 FT WO9621042-A2.  
 PN 11-JUL-1996.  
 PD 04-JAN-1996; U00461.  
 PR 04-JAN-1996; US-368706.  
 PA (UYBO-) UNIV. BOSTON.  
 PI DeLstein RA, Moreland RB;  
 DR WPI: 96-334017/33.  
 DR P-PSDB; W03130.  
 PT Detecting a metastatic disorder in a fixed biological sample -  
 PT comprises reverse transcription PCR of sample nucleic acid and  
 PT detection of metastatic, esp. prostate, tissue-specific target  
 PT sequence  
 PS Disclosure: Fig 1: 60bp: English.  
 CC Prostate-specific antigen (PSA) cDNA (T35867) and genomic (T35868)  
 CC sequences were used to design primers (see also T35855-62)  
 CC used for the specific amplification of PSA nucleic acids in  
 CC fresh and archival tissue samples as a means of detecting  
 CC metastatic sequences. The sequences code for portions (W03130  
 CC and W03131, respectively) of PSA, a serine protease found in  
 CC normal and neoplastic prostate epithelium.  
 CC Sequence 1728 BP; 405 A; 491 C; 449 G; 383 T;  
 SQ

Query Match 98.7%; Score 312; DB 1; Length 1728;  
 Best Local Similarity 60.8%; Pred. No. 2.22e-15;  
 Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

DB 792 GAGGAGTCTTGACCCCAAGAACTGAGTGTGAGCCTCCATGTTATTCAGTAC 851  
 QY 1 GARGARTTYTNACCCNARAAATYTCARTGTGATGATTCATGATTCATGATGATG 60  
 DB 852 GTGTGTGCGCAAGT 865  
 QY 61 GTTGTGCGCAAGT 74

RESULT 7  
 ID T04864 standard; cDNA; 1729 BP.  
 AC T04864;  
 DT 16-JAN-1996 (first entry)  
 DE Prostate-specific antigen cDNA.  
 KW Prostate-specific antigen; prostate cancer; diagnosis; ss.  
 OS Homo sapiens.  
 FH Key  
 FT cds Location/Qualifiers  
 FT 378..1091  
 FT /\*tag= a  
 FT WO9528498-A1.  
 PN 26-OCT-1995.  
 PD 14-APR-1995; U004680.  
 PR 15-APR-1994; US-229391.  
 PA (UYCO) UNIV. COLUMBIA NEW YORK.  
 PI Butteyan R, Katz AE, Olsson CA, Rafto A;  
 DR WPI: 95-373812/48.

DR P-PSDB; R77098.  
 PT Method for enhancing prostate-specific antigen detection - provides  
 PT sensitive means to identify early stages of prostate cancer.  
 PS Disclosure: Page 67-69; 94pp: English.  
 CC A prostate-specific antigen (PSA) is detected by isolating mRNA from  
 CC a sample, producing cDNA, and subjecting the cDNA to RT-PCR using  
 CC primers specific for the human PSA coding region (given in T04864).  
 CC Sequence 1729 BP; 405 A; 496 C; 445 G; 383 T;  
 SQ

Query Match 98.7%; Score 312; DB 1; Length 1729;  
 Best Local Similarity 60.8%; Pred. No. 2.22e-15;  
 Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

DB 792 GAGGAGTCTTGACCCCAAGAACTGAGTGTGAGCCTCCATGTTATTCAGTAC 851  
 QY 1 GARGARTTYTNACCCNARAAATYTCARTGTGATGATTCATGATTCATGATGATG 60  
 DB 852 GTGTGTGCGCAAGT 865  
 QY 61 GTTGTGCGCAAGT 74

RESULT 8  
 ID T90074 standard; cDNA; 246 BP.  
 AC T90074;  
 DT 14-APR-1998 (first entry)  
 DE Prostate specific antigen oligo-epitope encoding DNA.  
 KW Prostate specific antigen; PSA; epitope; vaccine; immune response;  
 KM Prostatic tumour cell; human leukocyte antigen; HLA; cytotoxic;  
 KM T lymphocyte; cancer; ds.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN WO9735021-A2.  
 PD 25-SEP-1997.  
 PR 19-MAR-1997; U00454.  
 PR 20-MAR-1996; US-618936.  
 PA (USSH) US SEC DEPT HEALTH.  
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 PI Schiom J, Tsang K, Zaremba S;  
 DR WPI: 97-480226/44.  
 PT Prostate specific antigen multiple epitope peptide - useful in  
 PT vaccine to produce immune response against prostate specific antigen  
 PT in prostatic tumour cell  
 PT Claim 23: Page 67; 85pp: English.  
 CC The present sequence encodes a prostate specific antigen oligo-epitope  
 CC peptide (PSA-OP) for use in a virus vector of the present invention. The  
 CC PSA-OP comprises more than 1 adjoined PSA-OP which generates a prostate  
 CC specific response in a portion of the human population. The PSA-OP can  
 CC be used to produce PSA specific cytotoxic T lymphocytes, in vivo or in  
 CC vitro, which can kill or prevent the establishment or growth of PSA  
 CC positive tumour cells, i.e. to prevent or treat prostatic cancer. The  
 CC products may be used to elicit a PSA specific immune response. When a  
 CC recombinant virus is used to elicit the response it is coadministered  
 CC with a carrier, a biological response modifier e.g. interleukin 2, 6 or  
 CC 12, interferon, tumour necrosis factor, granulocyte/macrophage colony  
 CC stimulating factor or cyclophosphamide and an adjuvant selected from  
 CC RIBI Detox, QS21, alum and incomplete Freund's adjuvant. PSA-OP  
 CC comprises multiple PSA epitopes, allowing an immune response to PSA in  
 CC a population of humans having more than 1 HLA class I molecule type to  
 CC be generated.  
 CC Sequence 246 BP; 56 A; 51 C; 57 G; 82 T;  
 SQ

Query Match 97.2%; Score 307; DB 1; Length 246;  
 Best Local Similarity 60.3%; Pred. No. 6.08e-15;  
 Matches 44; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

DB 130 AAGAGTCTTGACCCCAAGAACTGAGTGTGAGCCTCCATGTTATTCAGTAC 189  
 QY 2 ARGARTTYTNACCCNARAAATYTCARTGTGATGATTCATGATTCATGATGATG 61  
 DB 190 TGTGTGCGCAAGT 202  
 QY 62 TTTGTGCGCAAGT 74

FT	Intron	1848	3476	/+tag=	9
FT		/number=	2	3477	.3763
FT	exon	/+tag=	h		
FT		/number=	3	3764	.3906
FT	Intron	/+tag=	1		
FT		/number=	3	3907	.4043
FT	exon	/+tag=	j		
FT		/number=	4	4044	.5412
FT	Intron	/+tag=	k		
FT		/number=	4	5413	.5568
FT	exon	/+tag=	l		
FT		/number=	5		
PN		W09711172-A1.			
PD	27-MAR-1997				
PF	20-SEP-1996	U15123.			
PR	20-SEP-1995	US-004044.			
PA	(MORC-) WORCESTER FOUND BIOMEDICAL RES.				
PI	Zamecnik Pa;				
DR	WPI: 97-202881/18.				
DR	P-PSDB: W13649.				
PT	Treatment of benign prostatic hyperplasia or prostatic cancer				
PT	using an antisense oligonucleotide targeted against prostate				
PS	specific antigen or the probasin gene				
PS	Claim 2, Page 23-30; 49pp; English.				
CC	A novel method has been developed for treating a patient diagnosed as				
CC	having benign prostatic hyperplasia or a prostatic cancer. The method				
CC	comprises administering to the patient a therapeutic amount of a				
CC	composition comprising an antisense oligonucleotide which selectively				
CC	hybridises to prostate specific antigen (PSA) gene or mRNA or to a				
CC	probasin gene or mRNA sequence of the patient, where the antisense				
CC	oligonucleotide inhibits expression of the sequence. The present				
CC	sequence encodes human PSA. The antisense oligonucleotide can be used				
CC	to effectively inhibit the growth of (and for the killing of)				
CC	hyperplastic cells or cancerous cells of prostatic origin. Expression				
CC	of the PSA and probasin genes is specific to prostate cells so that the				
CC	antisense oligonucleotide can be administered systemically, making it				
CC	particularly useful in late stage prostatic cancer which has				
CC	metastasised, and in which the cells have become resistant to oestrogen				
CC	or anti-androgen therapy. The method can also be used in benign prostatic				
CC	hyperplasia, or early stage prostatic cancer, as an alternative for the				
CC	more radical procedures currently used, such as transurethral resection,				
CC	radical prostatectomy, or physical or chemical castration.				
SO	Sequence 5873 BP; 1186 A; 1778 C; 1503 G; 1406 T;				
Query Match	87.7%; Score 277; DB 1; Length 5873;				
Best Local Similarity	59.7%; Pred. No. 2,46e-12;				
Matches	40; Conservative 17; Mismatches 10; Indels 0; Gaps 0;				
Db	3907 TCTTGGCCCCCAAGAACTCAGTGTGACCTCCATGTTATTTCCAAATGACGTGTGG 3966				
OY	8 TTTTACCCCAAAATATTCARTGTGTGATGATTTTCAATGATGATGATG 67				
Db	3367 CGCAAGT 3973				
OY	68 CNCARGT 74				
RESULT	11				
ID	T90073 standard; cDNA; 90 BP.				
AC	T90073;				
DT	14-APR-1998 (first entry)				
DE	Prostate specific antigen oligo-epitope encoding DNA.				
KM	Prostate specific antigen; PSA; epitope; vaccine; immune response;				
KW	Prostatic tumor cell; human leukocyte antigen; HLA; cytotoxic;				
OS	T lymphocyte; cancer; ds.				
OS	Synthetic.				
OS	Homo sapiens.				

```

FH Key Location/Qualifiers
FT misc_difference 58..60
FT /tag= a
FT /transl_except= (pos:58..60,aa:lys)
PN MO9735021-A2.
PD 25-SEP-1997.
PF 19-MAR-1997: U04454.
PR 20-MAR-1996: US-618936.
PA (USSH ) US SEC DEPT HEALTH.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Schlom J, Tsang K, Zaremba S;
DR P-PSDB; W27387.
DR WPI: 97-480226/44.
PT Prostate specific antigen multiple epitope peptide - useful in
PT vaccine to produce immune response against prostate specific antigen
PI in prostatic tumor cell
PI Claim 21: Page 64; 85pp; English.
PS The present sequence encodes a prostate specific antigen oligo-epitope
CC peptide (PSA-OP) of the present invention. The PSA-OP comprises more
CC than 1 adjoined PSA-OP which generates a prostate specific response in
CC a portion of the human population. The PSA-OP can be used to produce
CC PSA specific cytotoxic T lymphocytes, in vivo or in vitro, which can
CC kill or prevent the establishment or growth of PSA positive tumor
CC cells, i.e. to prevent or treat prostatic cancer. The products may be
CC used to elicit a PSA specific immune response. When a recombinant virus
CC is used to elicit the response it is coadministered with a carrier, a
CC biological response modifier e.g. interleukin 2, 6 or 12, interferon,
CC tumor necrosis factor, granulocyte/macrophage-colony stimulating factor
CC or cyclophosphamide and an adjuvant selected from Ribi Detox, QS21,
CC alum and incomplete Freund's adjuvant. PSA-OP comprises multiple PSA
CC epitopes, allowing an immune response to PSA in a population of humans
CC having more than 1 HLA class I molecule type to be generated.
SQ Sequence 90 BP; 23 A; 24 C; 19 G; 24 T;

Query Match
Best Local Similarity 86.1%; Score 272; DB 1; Length 90;
Matches 40; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

DB 1 TTCTGACCCCAAGAACTTCAGTGTGAGCCTCCATGTCATATTCACATGAGCT 60
QY 7 TTYTANCCNCCNARARATNCARTGTGNGAYTNCATGTAHMSNAGAYGINTGT 66
DB 61 GCGCAGT 68
QY 67 GCMCAKGT 74

RESULT 12
ID V06602 standard; cDNA; 711 BP.
AC V06602:
DE Mature prostate-specific glandular kallikrein hk2 coding region.
DE Prostate-specific glandular kallikrein; hk2 protein; hK2 gene;
KW Prostate carcinoma; prostate cancer; benign prostatic hyperplasia;
KW antigen; diagnosis; human; ds.
OS Homo sapiens.
PN MO9802748-A1.
PD 22-JAN-1998.
PF 15-JUL-1997: U12322.
PR 15-JUL-1996: US-680868.
PA (HYBR-) HYBRITTECH INC.
PA (MAYO-) MAYO FOUNDATION.
PI Grauer L, Klee GG, Mikolajczyk SD, Saedi M, Tindall DJ,
PI Young CYF;
DR WPI: 98-120378/11.
DR P-PSDB; W45395.
PT Diagnostic methods using antibodies which bind prostate antigens -
PT useful for, e.g. monitoring treatment or progression of prostate
PT cancer
PS Example 1: Page 49-50; 100pp; English.
CC This cDNA comprises the coding region for human mature
CC prostate-specific glandular kallikrein hk2 (see W45395). cDNA
CC (see V06604) encoding the entire prepro-hk2 (pphk2, see W45397) was

```

```

CC synthesised from RNA of human benign prostate hyperplasia (BPH)
CC tissue by RT-PCR using hk2-specific primers (see V06605-06), and
CC inserted into vector pVL1393. Plasmid pPhk2/pVL1393 in E. coli
CC HB101 is deposited as ATCC 69614. The cDNA for the mature hk2
CC peptide can be utilised in the production of recombinant hk2 in
CC host cells. The invention relates to diagnostic methods employing
CC antibodies which bind to prostate antigen hk2 or pro-hk2 (see also
CC W45396) and which do not react significantly with prostate specific
CC antigen hk3 (see W45398). The methods are useful for the diagnosis
CC of prostate cancer and for monitoring the treatment and/or
CC progression, or the early detection, of prostate cancer in a male
CC who has BPH or a high grade prostatic neoplasia (HPGN) or a male
CC whose family members have or had BPH, HPGN or a prostate cancer.
SQ Sequence 711 BP; 153 A; 195 C; 210 G; 153 T;

Query Match
Best Local Similarity 47.1%; Score 181; DB 1; Length 711;
Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

DB 415 GAGGAGTTCGCGCCCGAGAGTTCAGTGTGAGCCTCCATGTCATGTCATGAC 474
QY 1 GAGGARTTTTNCNCCNARARATNCARTGTGNGAYTNCATGTAHMSNAGAY 60
DB 475 ATGTGTGC 482
QY 61 GTMTGYCC 68

RESULT 13
ID V70339 standard; cDNA; 711 BP.
AC V70339:
DE 09-FEB-1999 (first entry)
DE Prostate-specific glandular kallikrein protein hk2 encoding cDNA.
KW Prostate cancer; detection; hk2; hK1; hK3; pKh2; tissue kallikrein;
KW pphk2; prostate-specific glandular kallikrein protein; PSA; human;
KW prostate-specific antigen; ss.
OS Homo sapiens.
PI Slawin KM, Tindall DJ, Young CYF;
DR WPI: 98-594592/50.
DR P-PSDB; W83202.
DE Detection of human kallikrein 2 RNA - by reverse transcription and
DE amplification by PCR, for detecting, monitoring and staging of
DE prostate cancer
PS Example 1: Fig 2; 90pp; English.
CC The present invention describes a diagnostic method for detecting human
CC kallikrein 2 (hk2) DNA. The method comprises: (a) contacting DNA obtained
CC by reverse transcription (RT) of RNA from a human physiological sample
CC which comprises cells suspected of containing hk2 RNA with at least 2
CC oligonucleotides to amplify the DNA by PCR to yield amplified hk2 DNA,
CC where the conditions amplify the DNA obtained by RT of RNA from at least
CC one cell containing hk2 in a sample which comprises at least 107 to 109
CC cells; and (b) detecting the presence of the amplified hk2 DNA. The
CC method can be used for detecting, monitoring the progression of and
CC pathologically staging prostate cancer. The present sequence encodes
CC prostate-specific glandular kallikrein protein (hk2).
SQ Sequence 711 BP; 153 A; 195 C; 210 G; 153 T;

Query Match
Best Local Similarity 47.1%; Score 181; DB 1; Length 711;
Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

DB 415 GAGGAGTTCGCGCCCGAGAGTTCAGTGTGAGCCTCCATGTCATGTCATGAC 474
QY 1 GAGGARTTTTNCNCCNARARATNCARTGTGNGAYTNCATGTAHMSNAGAY 60

```



QY 1 GARGARTTYTNACNCNARARATNCARTGTCNGAYTTNCAYGTNATHWSNAYGAY 60  
 Db 475 ATGTGTGC 482  
 QY 61 GTNTGYGC 68

## RESULT 14

ID T05148 standard; cDNA; 754 BP.

AC T05148;

DT 04-MAR-1996 (first entry)

DE kallikrein hk2 cDNA.

KW kallikrein hk2; serine protease; prostate cancer; diagnosis; ds.

OS Homo sapiens.

PH key

FT CDS

PN MO930758-A1.

PD 16-NOV-1995

PF 09-MAY-1995; U06157.

PR 10-MAY-1994; US-241174.

PR 02-MAY-1995; US-427767.

PA (HYBR-) HYBRITTECH INC.

PA (MAYO-) MAYO FOUNDATION.

PI Saeedi MS, Tindall DJ, Young CYF;

DR WPI: 95-404123/51.

DR P-PSDB: R84670.

PT New isolated prostate-specific kallikrein polypeptide(s) - used to

PT develop prods. for use in assays for such polypeptide(s), partic.

PS Disclosure: Page 35-36; 61pp: English.

CC cDNA (T05148) coding for human hk2 kallikrein (R84670) is obt'd. by

CC PCR amplification of prepro-hk2 cDNA (T05147) using hk2-specific

CC primers (T05150, T05152). The cDNA is used for large-scale prodn.

CC In prokaryotic (E. coli) or eukaryotic (insect, mammalian) cells of

CC recombinant hk2, a prostate-specific glandular kallikrein

CC Sequence 754 BP; 164 A; 211 C; 215 G; 164 T;

SO

Query Match 57.3%; Score 181; DB 1; Length 754;  
 Best Local Similarity 47.1%; Pred. No. 2.57e-04;  
 Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

Db 424 GAGGAGTTCCTGCGCCCGAGAGTCTTCAGTGTGAGCCCTCCCTGTCATGAC 483  
 QY 1 GARGARTTYTNACNCNARARATNCARTGTCNGAYTTNCAYGTNATHWSNAYGAY 60  
 Db 484 ATGTGTGC 491  
 QY 61 GTNTGYGC 68

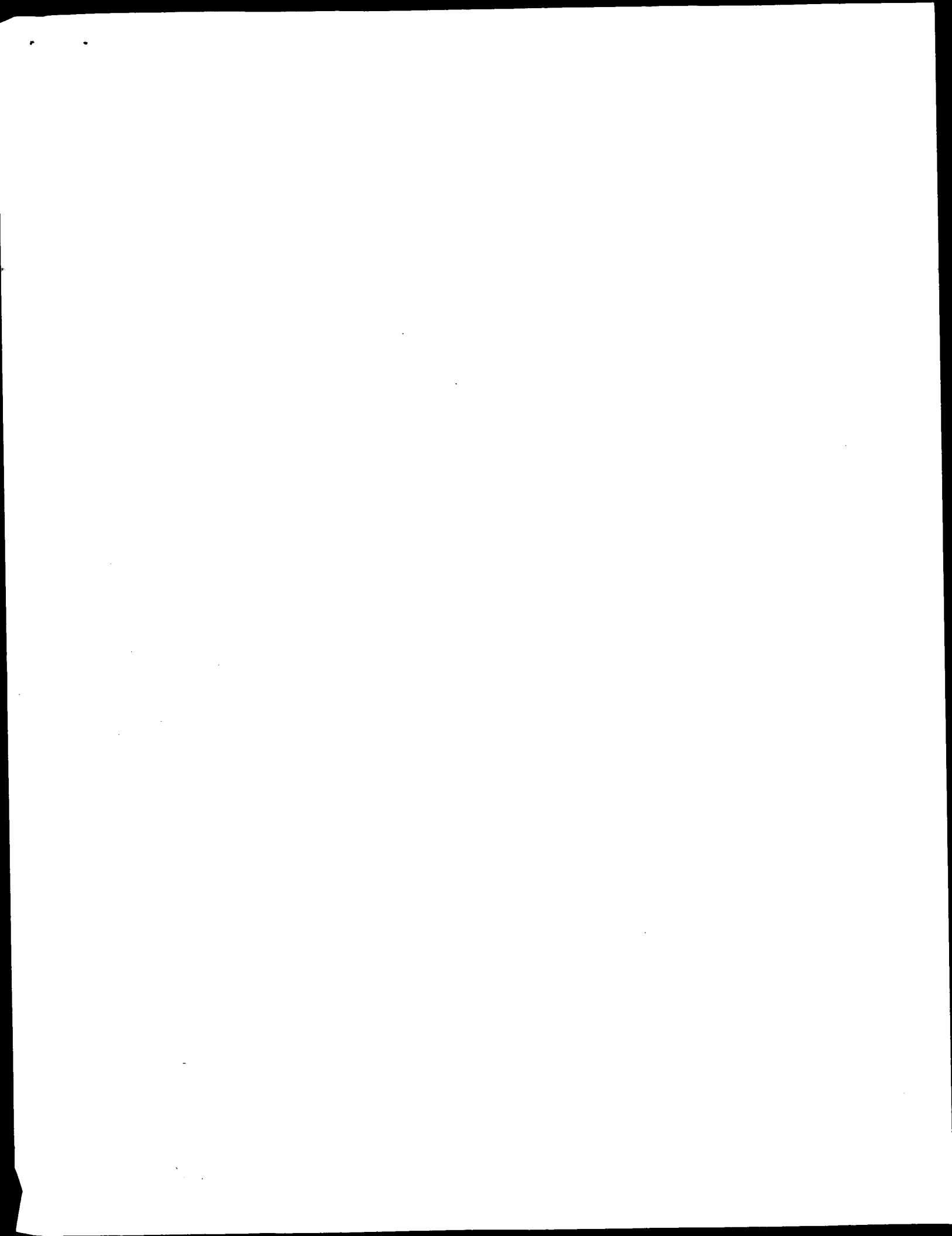
RESULT 15  
 ID V70341 standard; cDNA; 766 BP.  
 AC V70341;  
 DT 09-FEB-1999 (first entry)  
 DE Prostate-specific glandular kallikrein protein phk2 encoding cDNA.  
 KW prostate cancer; detection; hk2; hk1; hk3; phk2; tissue kallikrein;  
 KW phk2; prostate-specific glandular kallikrein protein; PSA; human;  
 OS Homo sapiens.  
 PH key  
 FT CDS  
 PN MO9846795-A1.  
 PD 22-OCT-1998.  
 PF 09-APR-1998; U07027.  
 PR 11-APR-1997; US-843076.  
 PA (BAYU) BAYLOR COLLEGE MEDICINE.  
 PA (MAYO-) MAYO FOUNDATION.  
 PI Slawin KM, Tindall DJ, Young CYF;  
 DR WPI: 98-594592/50.  
 DR P-PSDB: W83204.  
 PT Detection of human kallikrein 2 RNA - by reverse transcription and

PT amplification by PCR, for detecting, monitoring and staging of  
 PT prostate cancer  
 PS Example 1; Fig 2; 90pp: English.  
 CC The present invention describes a diagnostic method for detecting human  
 CC kallikrein 2 (hk2) DNA. The method comprises: (a) contacting DNA obtained  
 CC by reverse transcription (RT) of RNA from a human physiological sample  
 CC which comprises cells suspected of containing hk2 RNA with at least 2  
 CC oligonucleotides to amplify the DNA by PCR to yield amplified hk2 DNA,  
 CC where the conditions amplify the DNA obtained by RT of RNA from at least  
 CC one cell containing hk2 in a sample which comprises at least 107 to 109  
 CC cells; and (b) detecting the presence of the amplified hk2 DNA. The  
 CC method can be used for detecting, monitoring the progression of and  
 CC pathologically staging prostate cancer. The present sequence encodes  
 CC Prostate-specific glandular kallikrein protein (phk2).  
 SO Sequence 766 BP; 162 A; 218 C; 220 G; 166 T;

Query Match 57.3%; Score 181; DB 1; Length 766;  
 Best Local Similarity 47.1%; Pred. No. 2.57e-04;  
 Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

Db 436 GAGGAGTTCCTGCGCCCGAGAGTCTTCAGTGTGAGCCCTCCCTGTCATGAC 495  
 QY 1 GARGARTTYTNACNCNARARATNCARTGTCNGAYTTNCAYGTNATHWSNAYGAY 60  
 Db 496 ATGTGTGC 503  
 QY 61 GTNTGYGC 68

Search completed: Fri Jun 23 17:11:24 2000  
 Job time : 32 secs.



\*\*\*\*\*  
MUSEUM  
\*\*\*\*\*  
(TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mperch\_tpm n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:11:42 2000; Maspar time 15.09 Seconds  
645.987 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTNACNCNNA.....AYGATNTGYGCNARGTN 75  
Comp: CTCTCAARARANTGNGNTT.....TCTRCNACNCGNGTTCAN

Scoring table: TABLE bkttranslated2  
Gap 40

Mmatch STD : Dbase 0; Query 0

Searched: 230473 seqs, 64993002 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-1sued  
1:5A\_COMB 2:5B\_COMB 3:5C\_COMB 4:5D\_COMB 5:6\_COMB  
6:PCT\_COMB 7:backfile1

Statistics: Mean 37.556; Variance 145.681; scale 0.258

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	312	98.7	992	4	US-08-764-	Sequence 13, Applicat	1.18e-18
2	312	98.7	992	1	US-08-358-	Sequence 13, Applicat	1.18e-18
3	312	98.7	1462	1	US-08-358-	Sequence 14, Applicat	1.18e-18
4	312	98.7	1462	1	US-08-764-	Sequence 14, Applicat	1.18e-18
5	312	98.7	1729	4	US-08-718-	Sequence 1, Applicatio	1.18e-18
6	312	98.7	1729	4	US-08-844-	Sequence 1, Applicatio	1.18e-18
7	181	57.3	760	5	US-08-768-	Sequence 7, Applicatio	5.30e-06
8	181	57.3	760	5	PCT-US95-0	Sequence 7, Applicatio	5.30e-06
9	181	57.3	766	6	PCT-US95-0	Sequence 9, Applicatio	5.30e-06
10	181	57.3	766	6	US-08-768-	Sequence 9, Applicatio	5.30e-06
11	181	57.3	832	5	US-08-768-	Sequence 9, Applicatio	5.30e-06
12	181	57.3	832	5	PCT-US95-0	Sequence 5, Applicatio	5.30e-06
13	181	57.3	832	5	US-08-768-	Sequence 5, Applicatio	5.30e-06
14	160	50.6	60	4	US-08-832-	Sequence 8, Applicatio	4.20e-04
15	120	38.0	2815	4	US-08-940-	Sequence 1, Applicatio	1.13e+00
16	120	38.0	2815	2	US-08-230-	Sequence 1, Applicatio	1.13e+00
17	120	38.0	2815	2	US-08-619-	Sequence 1, Applicatio	1.13e+00
18	110	34.8	396	1	US-08-470-	Sequence 43, Applicati	7.23e+00

19	110	34.8	396	1	US-08-470-	Sequence 121, Applicat	7.23e+00
20	110	34.8	1534	1 <th>US-07-973-</th> <th>Sequence 2, Applicatio</th> <th>7.23e+00</th>	US-07-973-	Sequence 2, Applicatio	7.23e+00
21	107	33.9	580	1 <th>US-08-078-</th> <th>Sequence 1, Applicatio</th> <th>1.25e+01</th>	US-08-078-	Sequence 1, Applicatio	1.25e+01
22	107	33.9	8067	2 <th>US-08-486-</th> <th>Sequence 8, Applicatio</th> <th>1.25e+01</th>	US-08-486-	Sequence 8, Applicatio	1.25e+01
23	107	33.9	8067	1 <th>US-08-117-</th> <th>Sequence 8, Applicatio</th> <th>1.25e+01</th>	US-08-117-	Sequence 8, Applicatio	1.25e+01
24	105	33.2	5502	5 <th>US-08-836-</th> <th>Sequence 1, Applicatio</th> <th>1.78e+01</th>	US-08-836-	Sequence 1, Applicatio	1.78e+01
25	102	32.3	702	1 <th>US-08-518-</th> <th>Sequence 1, Applicatio</th> <th>3.04e+01</th>	US-08-518-	Sequence 1, Applicatio	3.04e+01
26	102	32.3	1209	6 <th>PCT-US95-0</th> <th>Sequence 5, Applicatio</th> <th>3.04e+01</th>	PCT-US95-0	Sequence 5, Applicatio	3.04e+01
27	102	32.3	1213	6 <th>US-08-231-</th> <th>Sequence 10, Applicati</th> <th>3.04e+01</th>	US-08-231-	Sequence 10, Applicati	3.04e+01
28	102	32.3	2540	4 <th>US-08-486-</th> <th>Sequence 15, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 15, Applicati	3.04e+01
29	102	32.3	2540	4 <th>US-08-486-</th> <th>Sequence 15, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 15, Applicati	3.04e+01
30	102	32.3	3007	5 <th>US-08-486-</th> <th>Sequence 15, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 15, Applicati	3.04e+01
31	102	32.3	3070	5 <th>US-08-486-</th> <th>Sequence 15, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 15, Applicati	3.04e+01
32	102	32.3	3070	5 <th>US-08-486-</th> <th>Sequence 15, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 15, Applicati	3.04e+01
33	102	32.3	3070	5 <th>US-08-486-</th> <th>Sequence 15, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 15, Applicati	3.04e+01
34	102	32.3	3211	5 <th>US-08-486-</th> <th>Sequence 15, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 15, Applicati	3.04e+01
35	102	32.3	3211	4 <th>US-08-486-</th> <th>Sequence 21, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 21, Applicati	3.04e+01
36	102	32.3	3274	4 <th>US-08-486-</th> <th>Sequence 21, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 21, Applicati	3.04e+01
37	102	32.3	3731	5 <th>US-08-486-</th> <th>Sequence 31, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 31, Applicati	3.04e+01
38	102	32.3	3731	5 <th>US-08-486-</th> <th>Sequence 31, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 31, Applicati	3.04e+01
39	102	32.3	3935	3 <th>US-08-231-</th> <th>Sequence 19, Applicati</th> <th>3.04e+01</th>	US-08-231-	Sequence 19, Applicati	3.04e+01
40	102	32.3	4094	5 <th>US-08-480-</th> <th>Sequence 25, Applicati</th> <th>3.04e+01</th>	US-08-480-	Sequence 25, Applicati	3.04e+01
41	102	32.3	4157	5 <th>US-08-480-</th> <th>Sequence 25, Applicati</th> <th>3.04e+01</th>	US-08-480-	Sequence 25, Applicati	3.04e+01
42	102	32.3	4157	4 <th>US-08-486-</th> <th>Sequence 25, Applicati</th> <th>3.04e+01</th>	US-08-486-	Sequence 25, Applicati	3.04e+01
43	102	32.3	4234	2 <th>US-08-805-</th> <th>Sequence 1, Applicatio</th> <th>3.04e+01</th>	US-08-805-	Sequence 1, Applicatio	3.04e+01
44	102	32.3	4234	1 <th>US-08-486-</th> <th>Sequence 1, Applicatio</th> <th>3.04e+01</th>	US-08-486-	Sequence 1, Applicatio	3.04e+01
45	102	32.3	4361	3 <th>US-08-231-</th> <th>Sequence 23, Applicati</th> <th>3.04e+01</th>	US-08-231-	Sequence 23, Applicati	3.04e+01

## ALIGNMENTS

RESULT 1  
ID US-08-764-527A-13 STANDARD; DNA; UNC; 992 BP.  
AC xxxxxx

DE Sequence 13, Application US/08764527A  
CC Sequence 13, Application US/08764527A  
CC Patent No. 5939258  
CC GENERAL INFORMATION:

CC APPLICANT: Cioce, Carlo  
CC APPLICANT: Gomella, Leonard  
CC APPLICANT: Mulholland, S. Grant  
CC APPLICANT: Moreno, Jose  
CC APPLICANT: Fischer, Rainer  
CC TITLE OF INVENTION: Methods of Detecting Micrometastasis of  
CC NUMBER OF SEQUENCES: 14  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESS: Woodcock Washburn Kurtz Mackiewicz &  
CC STREET: One Liberty Place 46th. Floor  
CC CITY: Philadelphia  
CC STATE: PA  
CC ZIP: 19103

CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/764,527A  
CC FILING DATE:  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/358,782  
CC FILING DATE: 15-DEC-1994  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Beardell, Lori Y.  
CC REGISTRATION NUMBER: 34,293  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 215-568-3100  
CC TELEFAX: 215-568-3439  
CC INFORMATION FOR SEQ ID NO: 13:



```

ID  US-08-764-527A-14 STANDARD; DNA: UNC; 1462 BP.
AC  xxxxxx
DT
DE  Sequence 14, Application US/08764527A
CC  Sequence 14, Application US/08764527A
CC  Patent No. 5939258
CC  GENERAL INFORMATION:
CC  APPLICANT: Croce, Carlo
CC  APPLICANT: Gomella, Leonard
CC  APPLICANT: Mulholland, S. Grant
CC  APPLICANT: Moreno, Jose
CC  APPLICANT: Fischer, Rainer
CC  TITLE OF INVENTION: Methods of Detecting Micrometastasis of
CC  TITLE OF INVENTION: Prostate
CC  NUMBER OF SEQUENCES: 14
CC  CORRESPONDENCE ADDRESS:
CC  ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
CC  STREET: One Liberty Place 46th. Floor
CC  CITY: Philadelphia
CC  STATE: PA
CC  ZIP: 19103
CC  COMPUTER READABLE FORM:
CC  MEDIUM TYPE: Floppy disk
CC  COMPUTER: IBM PC compatible
CC  OPERATING SYSTEM: PC-DOS/MS-DOS
CC  SOFTWARE: PatentIn Release #1.0, Version #1.25
CC  CURRENT APPLICATION DATA:
CC  APPLICATION NUMBER: US/08/764,527A
CC  FILING DATE:
CC  CLASSIFICATION:
CC  PRIOR APPLICATION DATA:
CC  APPLICATION NUMBER: US/08/358,782
CC  FILING DATE: 15-DEC-1994
CC  ATTORNEY/AGENT INFORMATION:
CC  NAME: Beardsell, Lori Y.
CC  REGISTRATION NUMBER: 34,293
CC  REFERENCE/DOCKET NUMBER: TJU-1327
CC  TELECOMMUNICATION INFORMATION:
CC  TELEPHONE: 215-568-3100
CC  TELEFAX: 215-568-3439
CC  INFORMATION FOR SEQ ID NO: 14:
CC  SEQUENCE CHARACTERISTICS:
CC  LENGTH: 1462 base pairs
CC  TYPE: nucleic acid
CC  STRANDEDNESS: single
CC  TOPOLOGY: linear
CC  MOLECULE TYPE: DNA (genomic)
CC  HYPOTHETICAL: NO
CC  ANTI-SENSE: NO
CC  SEQUENCE 1462 BP; 343 A; 374 C; 422 G; 323 T; 0 OTHER.
SO

Query Match          98.7%; Score 312; DB 4; Length 1462;
Best Local Similarity 60.8%; Pred.No.1.18e-18;
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db  511 GAGGAGTCTTGACCCCAAGAAGACTCAGGTGTGGACCTCCATGTTATTCGAATGAC 570
    |||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:
Oy  1 GARGRTTYTNNACCCCAARAAARATYTCARGTGTGTCNATYTNCAATGATNATHTMSNAAYGAY 60
    |||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:

Db  571 GTGTGTGGCAAGT 584
    |||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:
Oy  61 GTNTGYCGNCARGT 74

RESULT      5
ID  US-08-718-547-1 STANDARD; DNA: UNC; 1729 BP.
AC  xxxxxx
DT
DE  Sequence 1, Application US/08718547
CC  Sequence 1, Application US/08718547
CC  Patent No. 5876794
CC  GENERAL INFORMATION:
CC  APPLICANT: Katz, Aaron E.; Buttyan, Ralph; Rafto, Anthony;

```

```
CC APPLICANT: Olsson, Carl A.  
CC TITLE OF INVENTION: A Method For Molecular Staging Of  
CC TITLE OF INVENTION: Prostate Cancer  
CC NUMBER OF SEQUENCES: 4  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Cooper & Dunham LLP  
CC STREET: 1185 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: U.S.A.  
CC ZIP: 10036  
CC  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/718,547  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: White, John P.  
CC REGISTRATION NUMBER: 28,678  
CC REFERENCE/DOCKET NUMBER: 43677-A-PCT/JPW/MJG  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 278-0400  
CC TELEFAX: (212) 391-0525  
CC TELEX:  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1729 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC ORIGINAL SOURCE:  
CC ORGANISM: Homo sapiens  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 378..1088  
CC SEQUENCE 1729 BP: 405 A; 492 C; 449 G; 383 T; 0 OTHER.  
SQ  
  
Cc Query Match 98.7%; Score 312; DB 4; Length 1729;  
Cc Best Local Similarity 60.8%; Pred.No.1.18e-18;  
Cc Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;  
  
Db 792 GAGGAGTCTTGACCCCCAAGAACCCTCAGTGTGGACCTCCATGTTATTTCACATGAC 851  
Oy |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Oy 1 GARCGRTTYTNAACNCNCAARAARAYTCARTGYGTNGAYYTNCAYTNATHWSNAAYGAY 60  
  
Db 852 GTGTGTGCCCAAGT 865  
Oy |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Oy 61 GTNTGYGCNCARGT 74  
  
RESULT 6  
ID US-08-844-024-1 STANDARD; DNA; UNC; 1729 BP.  
AC xxxxxx  
DT  
DE Sequence 1, Application US/08844024  
CC Sequence 1, Application US/08844024  
CC Patent No. 5840494  
CC GENERAL INFORMATION:  
CC APPLICANT: Katz, Aaron E., et al.  
CC TITLE OF INVENTION: A Method For Molecular Staging Of  
CC TITLE OF INVENTION: Prostate Cancer  
CC NUMBER OF SEQUENCES: 4  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Cooper & Dunham  
CC STREET: 30 Rockefeller Plaza  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: U.S.A.
```



```
CC TELEFAX: 612-339-3061
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 760 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 7..720
SQ SEQUENCE 760 BP: 164 A; 212 C; 218 G; 166 T; 0 OTHER.
Dy Query Match 57.3%; Score 181; DB 6; Length 760;
Best Local Similarity 47.1%; Pred.No. 5.30e-06;
Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;
Db 424 GAGAGATTCTTGCCCCCAGGAGTCTTCAGTGTGAGACCTCATCTCCGTCCAATGAC 483
Oy 1 GARARATYYTNACNCCMARARAAYTNCARTGYGTNGAAYTYNCAYGTNATHWSNAAYGAY 60
Dy 484 ATGTGTGC 491
Oy 61 GTNTGYGC 68

RESULT 9
ID PCR-US95-06157-9 STANDARD; DNA; UNC; 766 BP.
AC xxxxxx
DT Sequence 9, Application PC/TUS9506157
DE Sequence 9, Application PC/TUS9506157
CC GENERAL INFORMATION:
CC APPLICANT: Mayo Foundation for Medical Education
CC APPLICANT: and Research
CC APPLICANT: Hybritech Incorporated
CC APPLICANT: Tindall, Donald J.
CC APPLICANT: Young, Charles V.F.
CC APPLICANT: Saedi, Mohammed S.
CC TITLE OF INVENTION: Recombinant HK2 Polypeptide
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Schwegman, Lundberg & Woessner, P.A.
CC STREET: 3500 IDS Center
CC CITY: Minneapolis
CC STATE: MN
CC COUNTRY: USA
CC ZIP: 55402
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/06157
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Raasch, Kevin W.
CC REGISTRATION NUMBER: 35,561
CC REFERENCE/DOCKET NUMBER: 150.148WO1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 612-339-0331
CC TELEFAX: 612-339-3061
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 766 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
```

```

CC LOCATION: 1..732
SQ SEQUENCE 766 BP: 162 A; 218 C; 219 G; 167 T; 0 OTHER.

Query Match 57.3%; Score 181; DB 6; Length 766;
Best Local Similarity 47.1%; Pred. No. 5,30e-06;
Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

Db 436 GAGAGAGTCTCGCCGCCAGAGAGTCTTCAGAGTGTGAGCCGCCATCTCCGTCCATGAC 495
||:||||:| ||:|:| ||:||||:| ||:|:| ||:|:|
OY 1 GARGARTTYTNNACNCCNAAARARYTNCARIGCTGTCNAYTTNCAYGTTNATHWSNAAATGAY 60

Db 436 ATGTGTGC 503
| ||:|
OY 61 GTNTGYGC 68

RESULT 10
ID US-08-768-859A-9 STANDARD; DNA; UNC; 766 BP.
AC xxxxxx
DT

DE Sequence 9, Application US/08768859A
CC Sequence 9, Application US/08768859A
CC Patent No. 6013471
CC GENERAL INFORMATION:
CC APPLICANT: Tindall, Donald J.
CC APPLICANT: Young, Charles Y.F.
CC APPLICANT: Saedi, Mohammed S.
CC TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides
CC NUMBER OF SEQUENCES: 21
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
CC STREET: P.O. Box 2938
CC CITY: Minneapolis
CC STATE: MN
CC COUNTRY: USA
CC ZIP: 55402
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentln Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/768,859A
CC FILING DATE: 17-DECEMBER-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Woessner, Warren D.
CC REGISTRATION NUMBER: 30,440
CC REFERENCE/DOCKET NUMBER: 545.002US3
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 612-339-0331
CC TELEFAX: 612-339-3061
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 766 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..732
CC SEQUENCE 766 BP: 162 A; 218 C; 220 G; 166 T; 0 OTHER.

SQ

Query Match 57.3%; Score 181; DB 5; Length 766;
Best Local Similarity 47.1%; Pred. No. 5,30e-06;
Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

Db 436 GAGAGAGTCTTGCCGCCAGAGAGTCTTCAGTGTGAGCCCTCCATCTCCGTCCATGAC 495
||:||||:| ||:|:| ||:||||:| ||:|:| ||:|:|
OY 1 GARGARTTYTNNACNCCNAAARARYTNCARIGCTGTCNAYTTNCAYGTTNATHWSNAAATGAY 60

Db 436 ATGTGTGC 503
| ||:|
OY 61 GTNTGYGC 68

```

QY 61 GINTGTGC 68

RESULT 11  
ID US-08-768-859A-5 STANDARD; DNA; UNC; 832 BP.  
AC xxxxxx

DE Sequence 5, Application US/08768859A  
CC Sequence 5, Application US/08768859A  
CC Patent No. 6013471  
CC GENERAL INFORMATION:  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeedi, Mohammed S.  
CC TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.  
CC STREET: P.O. Box 2938  
CC CITY: Minneapolis  
CC STATE: MN  
CC COUNTRY: USA  
CC ZIP: 55402

## COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/768,859A  
CC FILING DATE: 17-DECEMBER-1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Woessner, Warren D.  
CC REGISTRATION NUMBER: 30,440  
CC REFERENCE/DOCKET NUMBER: 545,002US3  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC TELEFAX: 612-339-3061  
CC INFORMATION FOR SEQ ID NO: 5:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 832 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: double  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 10..792  
CC SEQUENCE 832 BP; 169 A; 237 C; 242 G; 184 T; 0 OTHER.

QY 496 GAGAGTCTTGGCCCCAGAGCTTCAAGTGTGAGCCTCCATCCTGTCATGAC 555  
|||:||||:| | | : | | | : |||:| | : |||:| | : |||:| | :  
1 GAGGARTTYTNACNCNARARATNCARTGTGATGAYTNCAYTNATHTMSNAAYGAY 60  
Db 556 ATGTGTGC 563  
| ||:| |  
QY 61 GINTGTGC 68

Query Match 57.3%; Score 181; DB 5; Length 832;  
Best Local Similarity 47.1%; Pred. No. 5.30e-06;  
Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;  
Db 496 GAGAGTCTTGGCCCCAGAGCTTCAAGTGTGAGCCTCCATCCTGTCATGAC 555  
|||:||||:| | | : | | | : |||:| | : |||:| | : |||:| | :  
1 GAGGARTTYTNACNCNARARATNCARTGTGATGAYTNCAYTNATHTMSNAAYGAY 60  
QY 556 ATGTGTGC 563  
| ||:| |  
QY 61 GINTGTGC 68

RESULT 12  
PCT-US95-06157-5 STANDARD; DNA; UNC; 832 BP.  
AC xxxxxx  
DE Sequence 5, Application PC/TUS9506157  
CC Sequence 5, Application PC/TUS9506157  
CC GENERAL INFORMATION:  
CC APPLICANT: Mayo Foundation for Medical Education

CC APPLICANT: and Research  
CC APPLICANT: Hydriltech Incorporated  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeedi, Mohammed S.  
CC TITLE OF INVENTION: Recombinant HK2 Polypeptide  
CC NUMBER OF SEQUENCES: 18  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Schwegman, Lundberg & Woessner, P.A.  
CC STREET: 3500 IDS Center  
CC CITY: Minneapolis  
CC STATE: MN  
CC COUNTRY: USA  
CC ZIP: 55402

## COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/06157  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Raasch, Kevin W.  
CC REGISTRATION NUMBER: 35,561  
CC REFERENCE/DOCKET NUMBER: 150,148W01  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612-339-0331  
CC TELEFAX: 612-339-3061  
CC INFORMATION FOR SEQ ID NO: 5:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 832 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: double  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 10..792  
CC SEQUENCE 832 BP; 169 A; 237 C; 241 G; 185 T; 0 OTHER.

QY 496 GAGAGTCTTGGCCCCAGAGCTTCAAGTGTGAGCCTCCATCCTGTCATGAC 555  
|||:||||:| | | : | | | : |||:| | : |||:| | : |||:| | :  
1 GAGGARTTYTNACNCNARARATNCARTGTGATGAYTNCAYTNATHTMSNAAYGAY 60  
Db 556 ATGTGTGC 563  
| ||:| |  
QY 61 GINTGTGC 68

Query Match 57.3%; Score 181; DB 6; Length 832;  
Best Local Similarity 47.1%; Pred. No. 5.30e-06;  
Matches 32; Conservative 16; Mismatches 20; Indels 0; Gaps 0;  
Db 496 GAGAGTCTTGGCCCCAGAGCTTCAAGTGTGAGCCTCCATCCTGTCATGAC 555  
|||:||||:| | | : | | | : |||:| | : |||:| | : |||:| | :  
1 GAGGARTTYTNACNCNARARATNCARTGTGATGAYTNCAYTNATHTMSNAAYGAY 60  
QY 556 ATGTGTGC 563  
| ||:| |  
QY 61 GINTGTGC 68

RESULT 13  
ID US-08-768-859A-20 STANDARD; DNA; UNC; 832 BP.  
AC xxxxxx  
DE Sequence 20, Application US/08768859A  
CC Sequence 20, Application US/08768859A  
CC Patent No. 6013471  
CC GENERAL INFORMATION:  
CC APPLICANT: Tindall, Donald J.  
CC APPLICANT: Young, Charles Y.F.  
CC APPLICANT: Saeedi, Mohammed S.  
CC TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.  
CC STREET: P.O. Box 2938  
CC CITY: Minneapolis  
CC STATE: MN



RESULT 1  
 US-08-940-391-1 STANDARD; DNA; UNC; 2815 BP.  
 xxxxxx  
 Sequence 1, Application US/08940391  
 Sequence 1, Application US/08940391  
 Patent No. 5965373  
 GENERAL INFORMATION:  
 APPLICANT: Zimmermann, Rainer; Park, John E.;  
 APPLICANT: Rettig, Wolfgang; Old, Lloyd J.  
 TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION  
 PROTEIN ALPHA, AND USES THEREOF  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Felfe & Lynch  
 STREET: 805 Third Avenue  
 CITY: New York City  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10022  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Disette, 3.5 inch, 2.0 MB storage  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/940,391  
 FILING DATE: 01-OCT-1997  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/619,280  
 FILING DATE: 18-MARCH-1996  
 APPLICATION NUMBER: 08/230,491  
 FILING DATE: 20-APRIL-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Hanson, NO. 5965373man D.  
 REGISTRATION NUMBER: 30,946  
 REFERENCE/DOCKET NUMBER: LUD 5330.1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 688-9200  
 TELEFAX: (212) 838-3884  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:



22-Jun-1999  
 A32297; A32423; S03604; S05468; S05467; A32546; S02239  
 A26757; C31567; S00232; A23937; G01551  
 A33297

## #authors

#journal  
#journal  
#title  
#cross\_references  
#accession

Roelofs, J.F.C.; Vlastuin, K.J.; van der Korput, J.A.G.M.;  
Romijn, J.C.; Trapman, J.  
Biochem. Biophys. Res. Commun. (1989) 159:95-102  
Characterization of the prostate-specific antigen gene: a  
novel human kallikrein-like gene.  
Accession M0165891  
A32297

```
##molecule_type DNA
##molecule_subtype
```

```

#accession 1-261 #label RIE
#cross-references GB:M24543
REFERENCE
#authors Lundvall, A.
#journal Biochem. Biophys. Res. Commun. (1969) 161:1151-1159
#title Characterization of the gene for prostate-specific antigen, a
#cross-references M0ID:89302090
#accession A32423

```

## #molecu

```

#residues      1-261 ##label LUN
##cross-references GB:M27274: NID:g190552: PIDN:AAA60192.1: PID:g190553
REFERENCE      S03604

```

## #authors

```
#authors      Dipty, M.; Zhang, X.Y.; Richards, R.I.
#journal      Nucleic Acids Res. (1969) 17:2137
#title        Human prostate specific antigen (PSA) gene: structure and
               linkage to the kallikrein-like gene, hK1.
#cross-references MIM:89103632
#accession    503604
```

```
#Residuals      1-261 ##Label DIG
#Cross-correlation function
#Cross-correlation function
```

```

**residues      1-261 #label DIG
#cross-references EMBL:X13940
REFERENCE      S05468

```

#aucnors

```
#submission      Fittler, F.
#accession      submitted to the EMBL Data Library, May 1989
#sequence       S05468
#method comment
```

```

#residues      1-261 ##label KLI
#CROSS-referenced every 1000

```

REFERENCE  
#authors  
S05467  
Klobeck, H.G.; Combriato, G.; Schulz, P.; Arbusow, V.;  
Fittler, F.

## # Journal

Journal Nucleic Acids Res. (1989) 17:3981  
 title Genomic sequence of human prostate specific antigen (PSA)  
 cross-references W01D:89282407  
 accession 505467  
 molecule-type DNA  
 residues 1-29 ##label KL2  
 ##cross-references EMBL:XY010.0

Authors: P. Vihko, P. Henttu.

Journal: *Biochem. Biophys. Res. Commun.* (1989) 160:903-910  
 title: cDNA coding for the entire human prostate specific antigen  
 shows high homologies to the human tissue kallikrein genes  
 cross-references: MIM:602465  
 accession: A32546

#Residues 1-72, 'T', 74-85, '

```
261 #label HEN
##cross-references GB:M26663
#####
#####
```

STANDARD

Journal  
title  
Nucleic Acids Res. (1988) 16:6226  
Sequence of a cDNA clone encompassing the complete mature  
human prostate specific antigen (PSA) and an unspliced  
leader sequence.  
cross-references MUID:88289366

```
#accession      S02239
#molecule_type protein
```

```

#residues      17-63,'T',65-135,'M',137-261  ##label SCH
#cross-references EMBL:X07730
REFERENCE
NCCT57

```

## #authors

#authors Lundwall, A.; Lilje, H.  
#journal FEBS Lett. (1987) 214:317-322  
#title Molecular cloning of human prostate specific antigen cDNA  
#cross-references MIMD:87190978  
#accession A26757

```
##residuals 5-261 ##label LU2
REFERENCE 1000000
```

#ERECRL  
 #authors  
 #journal  
 #title  
 #cross-references  
 #accession  
 AV0144  
 Riegman, P.H.J.; Klaassen, P.; van der Korput, J.A.G.M.;  
 Romijn, J.C.; Trapman, J.  
 Biochem. Biophys. Res. Commun. (1988) 155:181-188  
 Molecular cloning and characterization of novel prostate  
 antigen CDNA's  
 accession MUID:88326297  
 C31567

## #Residues

```

#residues      5-261 ##label R12
#cross-references GB:M21895; NID:q189523; PID:AAA5995.1; PID:q189522
#note          the authors registered the codon GGC for residue 28 as
                Arg and TGG for residue 29 as Leu
REFERENCE      S00232

```

## #auto

#journal title	author	year	pages
Rickli, E.E. Eur. J. Biochem. (1987) 170:111-120 Isolation, characterization and amino-acid sequence of gamma-semi-inoprotein, a glycoprotein from human seminal plasma.		1987	11

```
#accession      time
#BOL00010      500232
```

#residues	25-261	#label	SC2
REFERENCE	A23937		

## #authors

#authors  
 #journal  
 #title  
 #cross-references

Lee, K.W.K.; Lee, P.J.; M'hamklu, T.; Chan, W.P.; Loor, R.  
 Proc. Natl. Acad. Sci. U.S.A. (1986) 83:3166-3170  
 Human prostate-specific antigen: structural and functional  
 similarity with serine proteases.  
 PMID:86205857

## #molecule\_

```

REFERENCE      G07735
##residues      25-93, 'T', 95-164, 'HL', 166, 'YDM', 169-174, 'Q', 176-261
##label      WAT

```

## #authors

```
#submission submitted to the EMBL Data Library, November 1994
#accession G01561
#status translated from GR/EMBL/DBD
#molecule_type mRNA
#residues 1-761 #label MOR
#cross-references EMBL:U17040; NID:g595945; PIDN:AAA56764.1
PENTONICS
PID:g595946
```

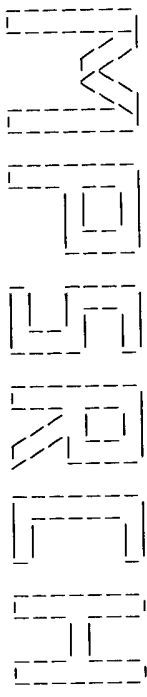
非非

```
#map_position 19q13.3-19q13.3
#introns
#classification
#superfamily trypsin; trypsin homology
#keywords glycoprotein; hydrolase; prostate; serine proteinase
#signature
```

18-24  
75-067

	product	prostatic	serine	protease	status
25-261	experimental	#label	MAV\		
25-253	#domain	trypsin	homology	#label	TRY\
65,120, 213	#active	site	HIS, Asp, Ser	status	predicted
UNIMMARY	#length	261	#molecular-weight	26741	#checksum-1355

```
Query Match      100.0%; Score 183; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.26e-28;
```



Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MSearch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 22 21:22:57 2000; MasPar time 5.79 Seconds

Tabular output not generated. 203,580 Million cell updates/sec

Title: >US-09-332-866-1

Description: (1-25) from US09332866.pep

Sequence: 1 EEFLPKKLCQVDLHVISNDVCAOV 25

Scoring table: PAM 150

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%

Database: Listing first 45 summaries

Database: p1r62

Statistics: Mean 28.844; Variance 40.415; scale 0.714

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	183	100.0	234	2	A31567	prostatic proantigen
2	183	100.0	261	2	A32297	prostate-specific ser
3	167	91.3	261	2	S35711	tissue kallikrein (EC
4	130	71.0	261	2	A29586	tissue kallikrein (EC
5	108	59.0	261	2	S45303	tissue kallikrein (EC
6	107	58.5	189	2	I52712	gene PSA protein - hu
7	96	55.5	257	2	S33772	tissue kallikrein (EC
8	91	49.7	262	1	KOHU	tissue kallikrein (EC
9	89	48.6	260	2	A37938	prostatic arginine es
10	89	48.6	261	2	S01971	tissue kallikrein (EC
11	87	47.5	250	2	S15685	kallikrein, glandular
12	87	47.5	261	2	A29745	tissue kallikrein (EC
13	84	45.9	261	2	A44284	tissue kallikrein (EC
14	83	45.4	244	2	KQRTN	tonin (EC 3.4.21.-) p
15	82	44.8	259	1	B23863	tissue kallikrein (EC
16	81	44.3	156	2	B33340	tissue kallikrein (EC
17	81	44.3	188	2	A31136	tissue kallikrein (EC
18	81	44.3	261	2	A34079	tissue kallikrein (EC
19	81	44.3	261	2	KQMS1	tissue kallikrein (EC
20	79	43.2	261	1	PC2013	tissue kallikrein (EC
21	78	42.6	246	2	JQ1471	trypsin (EC 3.4.21.4)
22	78	42.6	246	2	JQ1472	trypsin (EC 3.4.21.4)
23	78	42.6	246	2	JQ1472	trypsin (EC 3.4.21.4)

24	78	42.6	259	2	B31136	tissue kallikrein (EC
25	78	42.6	261	2	A25606	tissue kallikrein (EC
26	77	42.1	259	2	D23663	tissue kallikrein (EC
27	76	41.5	256	2	A05308	tissue kallikrein (EC
28	75	41.0	259	2	NGMSA	75 nerve growth facto
29	75	41.0	259	2	A29746	tissue kallikrein (EC
30	72	39.3	81	2	A18966	tissue kallikrein (EC
31	72	39.3	241	2	S39048	trypsin (EC 3.4.21.4)
32	71	38.8	237	1	KOPG	tissue kallikrein (EC
33	71	38.8	244	2	S72219	chymotrypsin B - Atla
34	69	37.7	240	2	S39047	trypsin (EC 3.4.21.4)
35	69	37.7	240	2	KQRTN	tissue kallikrein (EC
36	68	37.2	239	2	A27207	tissue kallikrein (EC
37	68	37.2	259	2	S68424	allergen der f iii pr
38	68	37.2	343	1	A57014	trypsin (EC 3.4.21.4)
39	67	36.6	250	2	T01779	trypsin (EC 3.4.21.4)
40	67	36.6	250	2	S31384	trypsin (EC 3.4.21.4)
41	67	36.6	253	2	A53968	serine proteinase SCC
42	67	36.6	261	2	A41020	tissue kallikrein (EC
43	67	36.6	261	2	JE0236	tissue kallikrein (EC
44	67	36.6	616	2	A30304	prolactin receptor 2
45	66	36.1	238	2	S31779	trypsin (EC 3.4.21.4)

## ALIGNMENTS

RESULT 1  
ENTRY A31567 #type fragment  
TITLE prostatic proantigen precursor (clone 525) - human (fragment)  
ORGANISM Homo sapiens #common\_name man  
DATE 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 22-Jun-1999

ACCESSIONS  
REFERENCE A31567  
#authors Riegman, P.H.J.; Klaassen, P.; van der Korp, J.A.G.M.; Romijn, J.C.; Trapman, J.

#journal Biochem. Biophys. Res. Commun. (1988) 155:181-188  
#title Molecular cloning and characterization of novel prostatic antigen cDNA's.  
#cross-references M01D:88326297

#accession A31567  
#molecule-type mRNA  
#residues 1-234 #label R1E  
#cross-references GB:M21896; NID:AA5996.1; PID:G189526  
##note the authors translated the codon GGC for residue 24 as Arg and TGG for residue 25 as Leu

CLASSIFICATION #superfamily trypsin; trypsin homology

FEATURE 1-13 #domain signal sequence (fragment) #status predicted

14-21 #label SIG  
21-234 #domain propeptide #status predicted #label PRO  
22-234 #domain trypsin homology #status atypical #label TRY  
61,116,209 #product prostatic antigen #status predicted #label MATV  
SUMMARY #length 234 #checksum 896

Query Match 100.0%; Score 183; DB 2; Length 234;  
Best Local Similarity 100.0%; Pred. No. 1,266-28;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 159 EEFLPKKLCQVDLHVISNDVCAOV 183  
QY 1 EEFLPKKLCQVDLHVISNDVCAOV 25

RESULT 2  
ENTRY A32297 #type complete  
TITLE prostate-specific serine proteinase (EC 3.4.21.-) precursor - human

ALTERNATE\_NAMES gamma-semioproprotein; prostatic antigen; tissue kallikrein (EC 3.4.21.35)-related proteinase, prostatic  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 08-Sep-1989 #sequence\_revision 08-Sep-1989 #text\_change

Best Local Similarity 100.0%; Pred. No. 1.40e-09; score 10/; DB 2; Length 189;

Query Match 58.5%; Score 107; DB 2; Length 189;  
Best Local Similarity 100.0%; Pred. No. 1.40e-09;

Matches 14: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 176 EEF1PKK1QCVLD 189  
1 EEF1PKK1QCVLD 14

QY 1 EEF1PKK1QCVLD 14

RESULT 7  
ENTRY S33772 #type complete  
TITLE tissue kallikrein (EC 3.4.21.35) precursor, renal -  
#formal name Macaca fascicularis #common\_name crab-eating  
macaque  
ORGANISM 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change  
22-Jun-1999

ACCESSIONS  
REFERENCE S33772  
#authors Lin, F.K.; Lin, C.H.; Chou, C.C.; Chen, K.; Lu, H.S.;  
Bacheller, W.; Herrera, C.; Jones, T.; Chao, J.; Chao, L.  
#journal Biochim. Biophys. Acta (1993) 1173:325-328  
#title Molecular cloning and sequence analysis of the monkey and  
human tissue kallikrein genes.  
#cross-references MUID:93305727  
#accession S33772  
#status preliminary  
#molecule\_type mRNA  
#residues 1-257 ##label LIN  
#cross-references GB:M31039; NID:9293140; PIDN:AAA36853.1; PID:9293141  
#note the authors translated the codon CCG for residue 17 as  
Ala, GCG for residue 18 as Arg, CCG for residue 54 as  
Arg, TCG for residue 71 as Met, and AAT for residue  
160 as His

CLASSIFICATION #superfamily trypsin; trypsin homology  
KEYWORDS hydrolase; serine proteinase

FEATURES  
25-249 #domain trypsin homology #label TRY  
62,116,209 #active-site His, Asp, Ser #status predicted  
SUMMARY #length 257 #molecular\_weight 28237 #checksum 7941

Query Match 52.5%; Score 96; DB 2; Length 257;  
Best Local Similarity 56.5%; Pred. No. 4.47e-07;  
Matches 13; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Db 159 ENFSFDD1QCVLDLIPNDECA 181  
1 EEF1PKK1QCVLDLIPNDECA 23

QY 1 EEF1PKK1QCVLDLIPNDECA 23

RESULT 8  
ENTRY KOHU #type complete  
TITLE tissue kallikrein (EC 3.4.21.35) precursor - human  
ALTERNATE\_NAMES kininogenin; pancreatic kallikrein; salivary kallikrein;  
urinary kallikrein  
ORGANISM #formal name Homo sapiens #common\_name man  
DATE 30-Sep-1987 #sequence\_revision 30-Sep-1987 #ext\_change  
18-Jun-1999

ACCESSIONS  
REFERENCE A24696; A28678; A23587; S05642; JX0040; A04628; A60248;  
#cross-references GB:M25629; NID:9186652; PIDN:AAA36136.1; PID:9186653  
A28678  
#authors Evans, B.A.; Yun, Z.X.; Close, J.A.; Tregear, G.W.; Kitamura,  
N.; Nakanishi, S.; Callen, D.F.; Baker, E.; Hyland, V.J.;  
Sutherland, G.R.; Richards, R.I.  
#title Nucleotide sequence of cloned cDNA for human pancreatic  
kallikrein.  
#cross-references MUID:86131629  
#accession A24696  
#molecule\_type mRNA  
#residues 1-262 ##label FUK  
#cross-references GB:M25629; NID:9186652; PIDN:AAA36136.1; PID:9186653  
A28678  
#authors Evans, B.A.; Yun, Z.X.; Close, J.A.; Tregear, G.W.; Kitamura,  
N.; Nakanishi, S.; Callen, D.F.; Baker, E.; Hyland, V.J.;  
Sutherland, G.R.; Richards, R.I.

REFERENCE  
#journal Biochemistry (1988) 27:3124-3129  
#title Structure and chromosomal localization of the human renal  
kallikrein gene.  
#cross-references MUID:88269498  
#accession A28678  
#molecule\_type DNA  
#residues 1-185, 'K', 187-262 ##label EVA  
#cross-references GB:M33105; GB:M33106; GB:M33107; GB:M33108;  
GB:M33109; NID:9186649; PIDN:AAA59455.1; PID:9186651

REFERENCE  
#journal DNA (1985) 4:445-450  
#authors Baker, A.R.; Shine, J.  
#title Human kidney kallikrein: cDNA cloning and sequence analysis.  
#cross-references MUID:86155264  
#accession A23587  
#molecule\_type mRNA  
#residues 17-262 ##label BAK

REFERENCE  
#journal Angermann, A.; Bergmann, C.; Appelhans, H.  
#authors Blochem. J. (1989) 262:787-793  
#title Cloning and expression of human salivary gland kallikrein in  
Escherichia coli.  
#cross-references MUID:90073574  
#accession S05642  
#molecule\_type mRNA  
#residues 1-144, 'E', 146-185, 'K', 187-262 ##label ANG  
#cross-references EMBL:X13561; NID:934026; PIDN:CAA31912.1; PID:934027  
#note the authors translated the codon GGT for residue 226 as  
Glu

REFERENCE  
#journal JX0040  
#authors Takahashi, S.; Irie, A.; Miyake, Y.  
#journal J. Biochem. (1988) 104:22-29  
#title Primary structure of human urinary prokallikrein.  
#cross-references MUID:89125217  
#accession JX0040  
#molecule\_type protein  
#residues 18-262 ##label TAK

REFERENCE  
#journal A04628  
#authors Lottspeich, F.; Geiger, R.; Henschen, A.; Kutzbach, C.  
#journal Hoppe-Seyler's Z. Physiol. Chem. (1979) 360:1947-1950  
#title N-terminal amino acid sequence of human urinary kallikrein  
homology with other serine proteases.  
#cross-references MUID:80114126  
#accession A04628  
#molecule\_type protein  
#residues 25-55 ##label LOT

REFERENCE  
#journal A60248  
#authors Kellermann, J.; Lottspeich, F.; Geiger, R.; Deutmann, R.  
#journal Adv. Exp. Med. Biol. (1989) 247A:519-525  
#title Human urinary kallikrein: amino acid sequence and  
carbohydrate attachment sites.  
#cross-references MUID:90102100  
#accession A60248  
#molecule\_type protein  
#residues 25-262 ##label KEL

REFERENCE  
#journal JX0218  
#authors Lu, H.S.; Lin, F.K.; Chao, L.; Chao, J.  
#journal Int. J. Pept. Protein Res. (1989) 33:237-249  
#title Human urinary kallikrein: complete amino acid sequence and  
sites of glycosylation.  
#cross-references MUID:89326688  
#accession JX0218  
#molecule\_type protein  
#residues 25-185, 'K', 187-262 ##label LUH

REFERENCE  
#journal A60251  
#authors Terashima, H.; Atomi, Y.; Ohnishi, N.; Kuroda, A.; Morioka,  
Y.; Ikeita, M.; Aoki, K.; Kamada, M.; Kituki, K.; Moriya,  
H.  
#journal Adv. Exp. Med. Biol. (1989) 247B:177-182  
#title Characterization of human pancreatic kallikrein.  
#cross-references MUID:90119035  
#accession A60251



```

Best Local Similarity 50.0%: Pred. No. 1.55e-05;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 165 FQTPDDLCQCSIKLLPNXYC 184
      1 11 1111 : ::::11
QY 3 FLTPKKKLCQVDLHVISNDVC 22

RESULT 12
ENTRY S15685 #type fragment
TITLE kallikrein, glandular - multimammate rat (Mastomys
        natalensis) (fragment)
ORGANISM #formal_name Mastomys natalensis
DATE 09-Jun-1994 #sequence_revision 10-Nov-1995 #text_change
        22-Jun-1999
ACCESSIONS S15685
REFERENCE S15685
AUTHORS Bell, R.A.; Fahnestock, M.
SUBMISSION submitted to the EMBL Data Library, December 1989
DESCRIPTION Sequence of a cDNA clone coding for a glandular kallikrein
        from Mastomys natalensis.
#accession S15685
#status preliminary
#molecule_type mRNA
#residues 1-250 ##label BEL
#cross-references EMBL:X17351; NID:g55524; PIDN:CAA35231.1; PID:e32221;
        PID:g1334152
CLASSIFICATION #superfamily trypsin; trypsin homology
FEATURE #domain trypsin homology #label TRY\
22-250 #active_site His, Asp, Ser #status predicted
62,117,211 #length 250 #checksum 4256
SUMMARY

Query Match 47.5%: Score 87; DB 2: Length 250;
Best Local Similarity 47.8%: Pred. No. 4.17e-05;
Matches 11; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Db 161 EEYKFPDLCQCVYLKMPNICA 183
      11: 1 1111 : ::::11
QY 1 EEFLPKKLCQVDLHVISNDVCA 23

RESULT 13
ENTRY A29745 #type complete
TITLE tissue kallikrein (EC 3.4.21.35) mGK-9 precursor,
        submandibular - mouse
ALTERNATE_NAMES glandular kallikrein mGK-9; major epidermal growth
        factor-binding protein
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
        22-Jun-1999
ACCESSIONS C29746; A29745; A27120; I70015
REFERENCE A90522
AUTHORS Drinkwater, C.C.; Evans, B.A.; Richards, R.I.
JOURNAL Biochemistry (1987) 26:6750-6756
TITLE Mouse glandular kallikrein genes: identification and
        characterization of the genes encoding the epidermal growth
        factor binding proteins.
#cross-references MUID:88107594
#accession C29746
#molecule_type DNA
#residues 1-261 ##label DRI
#cross-references GB:M17985; NID:q193476; PIDN:AAA37681.1; PID:g387166
#experimental_source strain BAB/c. salivary gland
REFERENCE A29745
AUTHORS Blaber, M.; Isaacson, P.J.; Bradshaw, R.A.
JOURNAL Biochemistry (1987) 26:6742-6749
TITLE A complete cDNA sequence for the major epidermal growth
        factor binding protein in the male mouse submandibular
        gland.
#cross-references MUID:88107593
#accession A29745
#molecule_type mRNA

```







\*\*\*\*\*  
 WIREIMAGE  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MSPrch\_bp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Thu Jun 22 21:22:03 2000; Maspar time 3.84 Seconds  
 Tabular output not generated. 198.438 Million cell updates/sec

Title: >US-09-332-866-1  
 Description: (1-25) from US09332866.pep  
 Perfect Score: 183  
 Sequence: 1 EEFLTPKILQCVLHVISNDCAQV 25

Scoring table: PAM 150  
 Gap 15

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: swiss-prot38  
 1:swissprot

Statistics: Mean 29.602; Variance 37.132; scale 0.797

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB	ID	Description	Pred. No.
1	183	100.0	261	1	PROS_HUMAN PROSTATE SPECIFIC ANTI	4.77e-32
2	167	91.3	261	1	PROS_HUMAN PROSTATE SPECIFIC ANTI	2.12e-27
3	130	71.0	261	1	KLK2_HUMAN GLANDULAR KALLIKREIN 2	5.07e-17
4	96	52.5	257	1	KLK1_MACFA GLANDULAR KALLIKREIN 1	5.50e-08
5	95	51.9	258	1	KLK1_PAPHA GLANDULAR KALLIKREIN 1	6.15e-08
6	91	49.7	262	1	KLK1_HUMAN GLANDULAR KALLIKREIN 1	5.75e-07
7	91	49.7	263	1	KLK1_PAPHA GLANDULAR KALLIKREIN 1	5.75e-07
8	89	48.6	260	1	ESTR_PANNA ARGININE ESTERASE PREC	1.73e-06
9	89	48.6	261	1	KLK1_MOUSE GLANDULAR KALLIKREIN K	4.51e-05
10	84	45.9	261	1	KLK1_MOUSE GLANDULAR KALLIKREIN K	4.51e-05
11	83	45.4	244	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
12	82	44.8	259	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
13	81	44.3	188	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
14	81	44.3	261	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
15	81	44.3	261	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
16	79	43.2	261	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
17	78	42.6	246	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
18	78	42.6	246	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
19	78	42.6	246	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
20	78	42.6	246	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
21	77	42.1	259	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
22	76	41.5	256	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05
23	75	41.0	259	1	KLK3_MOUSE TONIN PRECURSOR (EC 3	7.48e-05

24	72	39.3	232	1	KLK_PIG GLANDULAR KALLIKREIN (	1.23e-02
25	72	39.3	241	1	TRYX_GADMO TRYPSINOGEN X PRECURSO	1.23e-02
26	71	38.8	245	1	TRYX_GADMO TRYPSINOGEN B (EC 3.4	2.01e-02
27	69	37.7	241	1	TRYX_GADMO TRYPSINOGEN B (EC 3.4	5.29e-02
28	69	37.7	241	1	TRYX_GADMO TRYPSINOGEN B (EC 3.4	5.29e-02
29	68	37.2	239	1	KLK2_MOUSE GLANDULAR KALLIKREIN,	8.51e-02
30	68	37.2	239	1	KLK2_MOUSE GLANDULAR KALLIKREIN,	8.51e-02
31	68	37.2	239	1	KLK2_MOUSE GLANDULAR KALLIKREIN,	8.51e-02
32	67	36.6	253	1	TRYX_GADMO TRYPSINOGEN B (EC 3.4	5.29e-02
33	67	36.6	253	1	TRYX_GADMO TRYPSINOGEN B (EC 3.4	5.29e-02
34	67	36.6	253	1	TRYX_GADMO TRYPSINOGEN B (EC 3.4	5.29e-02
35	67	36.6	253	1	TRYX_GADMO TRYPSINOGEN B (EC 3.4	5.29e-02
36	66	36.1	238	1	KLK2_MOUSE GLANDULAR KALLIKREIN,	8.51e-02
37	66	36.1	238	1	KLK2_MOUSE GLANDULAR KALLIKREIN,	8.51e-02
38	65	35.5	1634	1	DEP3_SALSA TRYPSINOGEN III PRECUR	2.18e-01
39	64	35.0	352	1	DEP3_SALSA TRYPSINOGEN III PRECUR	2.18e-01
40	64	35.0	352	1	DEP3_SALSA TRYPSINOGEN III PRECUR	2.18e-01
41	63	34.4	135	1	KLK2_MOUSE GLANDULAR KALLIKREIN,	8.51e-02
42	63	34.4	135	1	KLK2_MOUSE GLANDULAR KALLIKREIN,	8.51e-02
43	63	34.4	135	1	KLK2_MOUSE GLANDULAR KALLIKREIN,	8.51e-02
44	63	34.4	135	1	KLK2_MOUSE GLANDULAR KALLIKREIN,	8.51e-02
45	63	34.4	1309	1	KLK2_MOUSE GLANDULAR KALLIKREIN,	8.51e-02

## ALIGNMENTS

RESULT 1	STANDARD:	PRT:	261 AA.
ID	PROS_HUMAN		
AC	P07288;		
DT	01-APR-1988 (Rel. 07, Created)		
DI	01-JUL-1989 (Rel. 11, last sequence update)		
DE	15-JUL-1999 (Rel. 38, last annotation update)		
DE	PROSTATE SPECIFIC ANTIGEN PRECURSOR (EC 3.4.21.77) (PSA) (GAMMA-SEMINOPROTEIN) (SEMINOGLASE) (SEMININ) (P-30 ANTIGEN).		
GN	APs.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;		
CC	Eutheria; Primates; Catarrhini; Homiidae; Homo.		
LN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-PROSTATE;		
RX	MEDLINE: 89183632.		
RA	Digby M.R., Zhang X.Y., Richard R.I.;		
RT	"Human prostate specific antigen (PSA) gene: structure and linkage to the kallikrein-like gene, hK1.1";		
RL	Nucleic Acids Res. 17:2137-2137(1989).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE: 89282407.		
RA	Klebeck H.-G., Compiato G., Schulz P., Arbusow V., Fittler F.;		
RT	"Genomic sequence of human prostate specific antigen (PSA).";		
RL	Nucleic Acids Res. 17:3981-3981(1989).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-LEUKOCYTE;		
RX	MEDLINE: 89302090.		
RA	Lundwall A.;		
RT	"Characterization of the gene for prostate-specific antigen, a human glandular kallikrein.";		
RL	Biochem. Biophys. Res. Commun. 161:1151-1159(1989).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-PROSTATE;		
RX	MEDLINE: 87190978.		
RA	Lundwall A., Lilja H.;		
RT	"Molecular cloning of human prostate specific antigen cDNA.";		
RL	FEBS Lett. 214:317-322(1987).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-PROSTATE;		
RX	MEDLINE: 89246551.		
RA	Henttu P., Viikro P.;		
RT	"cDNA coding for the entire human prostate specific antigen shows		

high homologues to the human tissue kallikrein genes.";  
 RL Biochem. Biophys. Res. Commun. 160:903-910(1989).  
 [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PROSTATE;  
 RX MEDLINE: 89165891.  
 RA Riegmans P.H.J., Vlietstra R.J., van der Korput J.A.G.M., Romijn J.C.,  
 Trapman J., "Characterization of the prostate-specific antigen gene: a novel  
 human kallikrein-like gene.";  
 RT Biochem. Biophys. Res. Commun. 159:95-102(1989).  
 [7]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PROSTATE;  
 RA Wonne M.M., Moreno J.M., Mele C.M., Mulholland G.M., Gomella L.G.,  
 Submitted (Dec-1994) to the EMBL/Genbank/DBP databases.  
 [8]  
 RP SEQUENCE OF 17-261 FROM N.A.  
 RX MEDLINE: 88289366.  
 RA Schulz P., Stucka R., Feldmann H., Combratio G., Klobeck H.-G.,  
 Pittler F., "Sequence of a cDNA clone encompassing the complete mature human  
 prostate specific antigen (PSA) and an unsplited leader sequence.";  
 RT Nucleic Acids Res. 16:6226-6226(1988).  
 [9]  
 RP SEQUENCE OF 25-261.  
 RX MEDLINE: 86205857.  
 RA Watt K.W.K., Lee P.J., M'Jinkulu T., Chan W.P., Loo R.,  
 "Human prostate-specific antigen: structural and functional  
 similarity with serine proteases.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 83:3166-3170(1986).  
 [10]  
 RP SEQUENCE OF 25-261.  
 RX MEDLINE: 88082806.  
 RA Schaller J., Akiyama K., Tsuda R., Hara M., Marti T., Rickli E.E.,  
 "Isolation, characterization and amino-acid sequence of gamma-  
 semiothrin, a glycoprotein from human seminal plasma.";  
 RT Eur. J. Biochem. 170:111-120(1987).  
 [11]  
 RP 3D-STRUCTURE MODELING.  
 RX MEDLINE: 95218633.  
 RA Villoutreix B.O., Gettoff E.D., Griffen J.H.,  
 "A structural model for the prostate disease marker, human prostate-  
 specific antigen.";  
 RT Protein Sci. 3:2033-2044(1994).  
 [12]  
 RP 3D-STRUCTURE MODELING.  
 RX MEDLINE: 98427950.  
 RA Coombs G.S., Bergstrom R.C., Pellegrini J.L., Baker S.I., Navre M.,  
 Smith M.M., Tainer J.A., Madison E., Corey D.R.,  
 "Substrate specificity of prostate-specific antigen (PSA)."  
 RT Chem. Biol. 5:475-488(1998).  
 CC -1- FUNCTION: PRESUMABLY HYDROLYZE THE HIGH MOLECULAR MASS SEMINAL  
 VESICLE PROTEIN THUS LEADING TO THE LIQUEFACTION OF THE SEMINAL  
 COAGULUM.  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: TYR-1-XAA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ALSO KNOWN AS THE  
 TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X05332; CAA28947.1; -  
 DR EMBL: X07730; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL: M27274; AAA60192.1; -  
 DR EMBL: M26663; AAA58802.1; -  
 DR EMBL: U17040; AAA56764.1; -  
 DR EMBL: M24543; AAA60193.1; ALT\_SEQ.  
 DR PIR: S03604; S03604.  
 DR PIR: A26757; A26757.  
 DR PIR: A32297; A32297.  
 DR PIR: A32423; A32423.  
 DR PDB: 1PSA; 26-JAN-95.  
 DR PDB: 2PSA; 16-MAR-99.  
 DR MTM: 176820; -  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE: PS00135; TRYPsin\_SER; 1.  
 DR PFAM: PF00089; trypsin; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Antigen; Zymogen; Signal;  
 3D-structure.  
 FT SIGNAL 1 18  
 FT PROPEP 16 24  
 FT CHAIN 25 261  
 FT ACT\_SITE 65 65  
 FT ACT\_SITE 120 120  
 FT ACT\_SITE 213 213  
 FT CARBOHYD 69 69  
 FT DISULFID 31 173  
 FT DISULFID 50 66  
 FT DISULFID 152 219  
 FT DISULFID 184 198  
 FT DISULFID 209 234  
 FT CONFLICT 64 64  
 FT CONFLICT 73 73  
 FT CONFLICT 86 86  
 FT CONFLICT 94 94  
 FT CONFLICT 136 136  
 FT CONFLICT 165 168  
 FT CONFLICT 175 175  
 FT CONFLICT 175 175  
 FT CONFLICT 184 184  
 FT CONFLICT 260 260  
 SQ SEQUENCE 261 AA; 28741 MM; AE9EF32AF872141A CRC64;  
 Query Match 100.0%; Score 183; DB 1; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 4.77e-32;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 163 EEFLEPKKLCQVDLVISNDVCAQV 187  
 QY 1 EEFLEPKKLCQVDLVISNDVCAQV 25  
 ID 2  
 AC PROS.MACMU STANDARD. PRT: 261 AA.  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE PROSTATE SPECIFIC ANTIGEN PRECURSOR (EC 3.4.21.35) (PSA) (GAMMA-  
 SEMINOPROTEIN).  
 GN APS.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;  
 OC Macaca.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 93363642.  
 RA Gauthier E.R., Chapdelaine P., Tremblay R.R., Dube J.Y.,  
 "Characterization of rhesus monkey prostate specific antigen cDNA.";  
 RT Biochim. Biophys. Acta 1174:207-210(1993).  
 CC -1- FUNCTION: PSA PRESUMABLY HYDROLYZE THE HIGH MOLECULAR MASS  
 SEMINAL VESICLE PROTEIN THUS LEADING TO THE LIQUEFACTION OF

CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-|-XAA BONDS IN  
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
CC KALLIDIN (LISYL-L-BRDYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
CC MET-|-XAA OR LEU-|-XAA.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ALSO KNOWN AS THE  
CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL, M18156: -; NOT\_ANNOTATED\_CDS.  
DR EMBL, M18157: AAA74454.1; -;  
DR EMBL, S39329: AAD13816.1; -;  
DR PIR, A29586: A29586.  
DR HSSP, P07288: 1PFA.  
DR MIM, 147960: -;  
DR PRINTS, PR00722: CHYMOTRYPSIN.  
DR PROSITE, PS00134: TRYPSIN\_HIS. 1.  
DR PROSITE, PS00135: TRYPSIN\_SER. 1.  
DR PFM, PF00089: trypsin. 1.  
KW Hydrolyase, Serine protease, Kininogenase, Glycoprotein;  
KM Multigene family; zymogen; signal.  
FT SIGNAL 1  
FT PROPEP 19 24 PROBABLE.  
FT CHAIN 25 261 ACTIVATION PEPTIDE (PROBABLE).  
FT ACT\_SITE 65 65 GLANDULAR KALLIKREIN 2.  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM.  
FT DISULFID 31 173 CHARGE RELAY SYSTEM.  
FT DISULFID 50 66 BY SIMILARITY.  
FT DISULFID 152 219 BY SIMILARITY.  
FT DISULFID 184 198 BY SIMILARITY.  
FT DISULFID 209 234 BY SIMILARITY.  
FT CARBOHYD 102 102 POTENTIAL.  
SQ SEQUENCE 261 AA; 28671 MW; 9CFEFA1162FE42D CRC64;  
  
Query Match 71.0%; Score 130; DB 1; Length 261;  
Best Local Similarity 66.7%; Pred. NO.5.07e-17;  
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
  
Db 163 EEF1PRSLQCVSLHLISNDMCAR 186  
||| | : ||| | : ||| | : ||| | :  
QY 1 EEF1PKKLCQVDLHVISNDVCAQ 24  
  
RESULT 4  
ID KIK1\_MACFA STANDARD; PRT; 257 AA.  
AC Q07276;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE GLANDULAR KALLIKREIN 1 PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN)  
DE (KIDNEY/PANCREAS/SALIVARY GLAND KALLIKREIN).  
OS KIK1.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;  
OC Macaca.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 93305727.  
RA Lin F.K., Lin C.H., Chou C., Chen K., Lu H.S., Bacheller B.,  
RA Herrera C., Jones T., Chao J., Chao L.;  
RT Molecular cloning and sequence analysis of the monkey and human  
RT tissue kallikrein genes.";  
RL Biochim. Biophys. Acta 1173:325-338(1993).  
CC -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
CC IN KININOGEN TO BRINEN AND KININ.  
CC

```
CC      SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE INVOLVES HYDROLISTS OF
CC      KALLIDIN (LYSYL-BRDYKININ). FROM KINIINOGEN INVOLVES HYDROLISTS OF
CC      MET-1-XAA OR LEO-1-XAA.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC      TRYPSIN FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; LA3121; AAA73523.1; -.
DR      HSSP; P00752; ZKA1.
DR      PRINTS; PRO0722; CHYMOTRYPSIN.
DR      PROSITE; PS00134; TRYPsin_HIS_1.
DR      PROSITE; PS00135; TRYPsin_SER_1.
DR      Pfam; PF00089; trypsin; 1.
KW      Hydrolase; Serine protease; kinnogenase; Glycoprotein; Pancreas;
KW      zymogen; Signal.
FT      SIGNAL          1       18      BY SIMILARITY.
FT      PROPEP         19       24      ACTIVATION PEPTIDE (PROBABLE) .
FT      CHAIN          25       258    GLANDULAR KALLIKREIN 1.
FT      ACT_SITE       62        62    CHANGE RELAY SYSTEM (BY SIMILARITY) .
FT      ACT_SITE       117       117    CHARGE RELAY SYSTEM (BY SIMILARITY) .
FT      ACT_SITE       210       210    CHARGE RELAY SYSTEM (BY SIMILARITY) .
FT      DISULFID       31        170    BY SIMILARITY.
FT      DISULFID       47         63    BY SIMILARITY.
FT      DISULFID       149       216    BY SIMILARITY.
FT      DISULFID       181       195    BY SIMILARITY.
FT      DISULFID       206       231    BY SIMILARITY.
FT      CARBOHYD       90         90    POTENTIAL.
FT      CARBOHYD       99         99    POTENTIAL.
FT      CARBOHYD      101        101    POTENTIAL.
FT      CARBOHYD      105        105    POTENTIAL.
FT      CARBOHYD      161        161    POTENTIAL.
SQ      SEQUENCE      258 AA; 28339 MW; 8254BEBE84F36B4 CRC64;

Query Match           51.9%; Score 95; DB 1; Length 258;
Best Local Similarity 56.3%; Pred. No. 6,15e+08;
Matches 13; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

DB      160 ENFSYDDLQCVDLKITLPNDKCA 182
CQ      I:I|I||||::|||I
        1 EFLLPKRIQCVDLVHYSNDVCA 23

RESULT 6
ID      KRL1_HUMAN      STANDARD;             PR1;      262 AA.
AC      P06870;
DT      01-JAN-1988 (Rel. 06, Created)
DT      01-JAN-1988 (Rel. 06, Last sequence update)
DT      01-NOV-1995 (Rel. 32, Last annotation update)
DE      GLANDULAR KALLIKREIN 1 PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN)
DE      (KIDNEY/PANCREAS/SALIVARY GLAND KALLIKREIN).
GN      KRL1.
OS      Homo sapiens (Human) .
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC      Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=PANCREAS;
RX      MEDLINE; 86131629.
RA      Fukushima D., Kitamura N., Nakanishi S.;
RT      "Nucleotide sequence of cloned cDNA for human pancreatic kallikrein.";
RL      Biochemistry 24:8037-8043(1985).
[2]
RN      SEQUENCE FROM N.A.
RP      TISSUE=KIDNEY;
RC      MEDLINE; 88269498.
RA      Evans B.A., Yun Z.X., Close J.A., Tregear G.W., Kitamura N.,
```

RA Nakanishi S., Callen D.F., Baker E., Hyland V.J., Sutherland G.R.,  
 RA Richards R.I.;  
 RT "Structure and chromosomal localization of the human renal kallikrein  
 RT gene";  
 RL Biochemistry 27:3124-3129(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SALIVARY GLAND;  
 RX MEDLINE: 90073574.  
 RA Angermann A., Bergmann C., Appelhaus H.;  
 RT "Cloning and expression of human salivary-gland kallikrein in  
 RT Escherichia coli";  
 RL Biochem. J. 262:787-793(1989).  
 RN [4]  
 RP SEQUENCE OF 17-262 FROM N.A.  
 RC TISSUE-KIDNEY;  
 RX MEDLINE: 86135264.  
 RA Baker A.R., Shine J.;  
 RT "Human kidney kallikrein: cDNA cloning and sequence analysis";  
 RL DNA 4:445-450(1985).  
 RN [5]  
 RP SEQUENCE OF 25-262.  
 RC TISSUE-URINE;  
 RX MEDLINE: 89326688.  
 RA Lu H.S., Lin F.-K., Chao L., Chao J.;  
 RT "Human urinary kallikrein. Complete amino acid sequence and sites of  
 RT glycosylation";  
 RL Int. J. Pept. Protein Res. 33:237-249(1989).  
 RN [6]  
 RP SEQUENCE OF 25-262, AND CARBOHYDRATE-BINDING SITES.  
 RC TISSUE-URINE;  
 RX MEDLINE: 88203586.  
 RA Kellermann J., Lotsepelch F., Geiger R., Deutzmann R.;  
 RT "Human urinary kallikrein -- amino acid sequence and carbohydrate  
 RT attachment sites";  
 RL Protein Seq. Data Anal. 1:177-182(1988).  
 RN [7]  
 RP SEQUENCE OF 25-55.  
 RC TISSUE-URINE;  
 RX MEDLINE: 80114126.  
 RA Lotsepelch F., Geiger R., Henschen A., Kutzbach C.;  
 RT "N-Terminal amino acid sequence of human urinary kallikrein homology  
 RT with other serine proteases";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:1947-1950(1979).  
 RN [8]  
 RP SEQUENCE OF 28-47.  
 RC TISSUE-URINE;  
 RX MEDLINE: 86223893.  
 RA Takahashi S., Irie A., Katayama Y., Ito K., Miyake Y.;  
 RT "N-terminal amino acid sequence of human urinary prokallikrein";  
 RL J. Biochem. 99:989-992(1986).  
 CC -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
 CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
 CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
 CC MET-1-XAA OR LEU-1-XAA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M25629; AAA36136.1; -  
 CC DR EMBL: M33109; AAA59455.1; -  
 CC DR EMBL: M33105; AAA59455.1; JOINED  
 CC DR EMBL: M33106; AAA58455.1; JOINED  
 CC EMBL: M33107; AAA58455.1; JOINED.

DR EMBL: M33108; AAA59455.1; JOINED.  
 DR EMBL: X13561; CAA31912.1; -  
 DR PIR: A24696; KOHU.  
 DR HSSP: P00757; 1SGF.  
 DR MIM: 147910; -  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PROSITE: PS00134; TRYPSIN\_HIS. 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER. 1.  
 DR PFAM: PF00089; trypsin.1.  
 KW Hydrolase; Serine protease; Kininogenase; Glycoprotein;  
 KW Multigene family; zymogen; signal.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 262  
 FT ACT\_SITE 65 65  
 FT ACT\_SITE 120 120  
 FT ACT\_SITE 214 214  
 FT CARBOHYD 93 93  
 FT CARBOHYD 102 102  
 FT CARBOHYD 104 104  
 FT CARBOHYD 108 108  
 FT CARBOHYD 165 165  
 FT CARBOHYD 167 167  
 FT DISULFID 31 174  
 FT DISULFID 50 66  
 FT DISULFID 153 220  
 FT DISULFID 185 199  
 FT DISULFID 210 235  
 FT VARIANT 145 145  
 FT VARIANT 145 145  
 FT VARIANT 186 186  
 FT VARIANT 186 186  
 FT FT  
 SQ SEQUENCE 262 AA; 28889 MR; 772AD14A3A496C0 CRC64;  
 /FTid=VAR.006625.  
 E -> K.  
 Query Match 49.7%; Score 91; DB 1; Length 262;  
 Best Local Similarity 54.5%; Pred. No. 5.75e-07;  
 Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 DB 164 ENFSFPDDQCVDTKLIPNDEC 185  
 QY 1 EFLLPKKLQCVDLHVISNDVC 22  
 ID 7  
 AC KLR PRANA STANDARD: PRT: 263 AA.  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE GLANDULAR KALLIKREIN, RENAL PRECURSOR (EC 3.4.21.35) (TISSUE  
 DE KALLIKREIN).  
 OC Prameys natalensis (African soft-furred rat) (Mastomys natalensis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mastomys.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SALIVARY GLAND;  
 RX MEDLINE: 94226702.  
 RA Fahnestock M.;  
 RT "Characterization of kallikrein cDNAs from the African rodent  
 RT Mastomys";  
 RL DNA Cell Biol. 13:293-300(1994).  
 CC -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ.  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
 CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
 CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
 CC MET-1-XAA OR LEU-1-XAA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X17352; CA35232.1; -  
DR PIR: S15686; S15686.  
DR HSSP: P00757; 1SGF.  
DR PRINTS: PRO0722; CHYMOTRYPSIN.  
DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
DR PFM: PFO0089; trypsin; 1.  
KW Hydrolyase; Serine protease; Kininogenase; Glycoprotein;  
KW Multigene family; zymogen; signal.  
FT SIGNAL 1  
FT PROPEP 19 24 ACTIVATION PEPTIDE (PROBABLE).  
FT CHAIN 25 263 GLANDULAR KALLIKREIN, RENAL.  
FT ACT\_SITE 65 65 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 121 121 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 215 215 CHARGE RELAY SYSTEM.  
FT DISULFID 31 175 BY SIMILARITY.  
FT DISULFID 50 66 BY SIMILARITY.  
FT DISULFID 153 221 BY SIMILARITY.  
FT DISULFID 186 200 BY SIMILARITY.  
FT DISULFID 211 236 BY SIMILARITY.  
FT CARBOHYD 102 102 PROBABLE.  
SQ SEQUENCE 263 AA; 29130 MW; A8E023B800337D5 CRC64;

Query Match 49.7%; Score 91; DB 1; Length 263;  
Best Local Similarity 56.5%; Pred. No. 5,75e-07;  
Matches 13; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Db 165 EEPFYSHDLCCVLELISNVCVCA 187  
||| : |||| : ||||: ||||  
QY 1 EEPFLPKKLCQVDLHVISNDVCA 23

RESULT 8 STANDARD: PRT; 260 AA.  
AC P09362;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE ARGININE ESTERASE PRECURSOR (EC 3.4.21.35).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-PROSTATE;  
RX MEDLINE: 88211858.  
RA Chapdelaine P., Ho-Kim M.-A., Tremblay R.R., Dube J.Y.;  
RT "Nucleotide sequence of the androgen-dependent arginine esterase mRNA  
of canine prostate.";  
RL FEBS Lett. 232:187-192(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 91119675.  
RA Chapdelaine P., Gauthier E., Ho-Kim M.A., Bissonnette L.,  
RA Tremblay R.R., Dube J.Y.;  
RT "Characterization and expression of the prostatic arginine esterase  
RT gene, a canine glandular kallikrein.";  
RL DNA Cell Biol. 10:49-59(1991).  
RN [3]  
RP SEQUENCE OF 25-50 AND 108-145.  
RC TISSUE-PROSTATE;  
RX MEDLINE: 85004070.  
RA Lazure C., Leduc R., Seidah N.G., Chretien M., Dube J.Y.,  
RA Chapdelaine P., Frenette G., Paquin R., Tremblay R.R.;  
RT "The major androgen-dependent protease in dog prostate belongs to the  
RT kallikrein family: confirmation by partial amino acid sequencing.";  
RL FEBS Lett. 175:1-7(1984).

CC -1- FUNCTION: THIS SERINE PROTEASE IS FOUND IN DOG SEMINAL PLASMA,  
CC ITS EXACT PHYSIOLOGICAL FUNCTION IS NOT KNOWN.  
CC -1- INDUCTION: BY ANDROGENS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Y00751; CA68720.1; -  
DR EMBL: M63669; AAA30831.1; -  
DR PIR: A30981; A30981.  
DR PIR: A37938; A37938.  
DR PIR: S00613; S00613.  
DR HSSP: P00752; 2KAT.  
DR PRINTS: PRO0722; CHYMOTRYPSIN.  
DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
DR PROSITE: PS00135; TRYPsin\_SER; 1.  
DR PFM: PFO0089; trypsin; 1.  
KW Hydrolyase; Serine protease; Glycoprotein; zymogen; signal.  
FT SIGNAL 1  
FT PROPEP 2 24 ACTIVATION PEPTIDE.  
FT CHAIN 25 260 ARGININE ESTERASE.  
FT ACT\_SITE 65 65 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 119 119 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT CARBOHYD 79 79 PROBABLE.  
FT DISULFID 31 172 BY SIMILARITY.  
FT DISULFID 50 66 BY SIMILARITY.  
FT DISULFID 151 218 BY SIMILARITY.  
FT DISULFID 183 197 BY SIMILARITY.  
FT DISULFID 208 233 BY SIMILARITY.  
FT CONFLICT 56 N -> H (IN REF. 2).  
SQ SEQUENCE 260 AA; 28746 MW; 48768B6FE204775A CRC64;

Query Match 48.6%; Score 89; DB 1; Length 260;  
Best Local Similarity 52.0%; Pred. No. 1,73e-06;  
Matches 13; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 162 ETIFPGSLCCVCDLKLNNOCAYV 186  
||| : |||| : ||||: ||||  
QY 1 EEPFLPKKLCQVDLHVISNDVCA 25

RESULT 9 STANDARD: PRT; 261 AA.  
AC K1KA.MOUSE  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-FEB-2000 (Rel. 39, Last annotation update)  
DE GLANDULAR KALLIKREIN K11 PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN)  
(MCK-11).  
DE K1K11 OR K1K-11.  
GN K1K11 OR K1K-11.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C; TISSUE=LIVER;  
RX MEDLINE: 89083511.  
RA Drinkwater C.C., Richards R.I.;  
RT "Sequence of mck-11, a mouse glandular kallikrein gene.";  
RL Nucleic Acids Res. 16:10918-10918(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 85257431.  
RA Evans B.A., Richards R.I.;  
RT "Genes for the alpha and gamma subunits of mouse nerve growth factor



RT are contiguous.";  
 RL EMBO J. 4:133-138(1985).  
 RN [3]  
 RP SEQUENCE OF 16-54 AND 69-122 FROM N.A.  
 RA MEDLINE: 87250386.  
 RX Evans B.A., Drinkwater C.C., Richards R.I.;  
 RT "Mouse glandular kallikrein genes. Structure and partial sequence  
 analysis of the kallikrein gene locus.";  
 RL J. Biol. Chem. 262:8027-8034(1987).  
 CC -1- FUNCTION: GLANDULAR KALLIKREIN CLEAVE MET-LYS AND ARG-SER BONDS  
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
 CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
 CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
 CC MET-1-XAA OR LEU-1-XAA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X13215; CAA31604.3; JOINED.  
 DR EMBL: X13216; CAA31604.3; JOINED.  
 DR EMBL: X13217; CAA31604.3; JOINED.  
 DR EMBL: X13218; CAA31604.3; JOINED.  
 DR EMBL: M18590; AAA39352.1; ALT\_SEQ.  
 DR EMBL: M18610; AAA39353.1; -  
 DR PIR: S01971; S01971.  
 DR HSSP: P00757; 1SGF.  
 DR MGD: MGI:892023; KIKL1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE: PS00135; TRYPsin\_SER; 1.  
 DR PFAM: PF00089; trypsin; 1.  
 DR Hydroxylase: Serine protease; Kininogenase; Glycoprotein;  
 DR Multigene family: zymogen; Signal.  
 DR FT SIGNAL 1 18  
 DR FT PROPEP 19 24  
 DR FT CHAIN 25 261  
 DR FT ACT\_SITE 65 65  
 DR FT ACT\_SITE 120 120  
 DR FT ACT\_SITE 213 213  
 DR FT DISULFID 31 173  
 DR FT DISULFID 50 66  
 DR FT DISULFID 152 219  
 DR FT DISULFID 198 198  
 DR FT DISULFID 209 234  
 DR FT CARBOHYD 102 102  
 DR SEQUENCE 261 AA; 28727 MW; 8D22EF8D835EA7B CAC64;  
 SO  
 Query Match 48.6%; Score 89; DB 1; Length 261;  
 Best Local Similarity 50.0%; Pred. No. 1.73e-06;  
 Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE-SALIVARY GLAND;  
 RX MEDLINE: 88107594.  
 RA Drinkwater C.C., Evans B.A., Richards R.I.;  
 RT "Mouse glandular kallikrein genes: identification and  
 RT characterization of the genes encoding the epidermal growth factor  
 RT binding proteins.";  
 RL Biochemistry 26:6750-6756(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 88107593.  
 RA Blaber M., Isackson P.J., Bradshaw R.A.;  
 RT "A complete cDNA sequence for the major epidermal growth factor  
 RT binding protein in the male mouse submandibular gland.";  
 RL Biochemistry 26:6742-6749(1987).  
 RN [3]  
 RP SEQUENCE OF 16-54 AND 70-122 FROM N.A.  
 RC STRAIN-BALB/C; TISSUE-LIVER;  
 RX MEDLINE: 87250386.  
 RA Evans B.A., Drinkwater C.C., Richards R.I.;  
 RT "Mouse glandular kallikrein genes. Structure and partial sequence  
 RT analysis of the kallikrein gene locus.";  
 RL J. Biol. Chem. 262:8027-8034(1987).  
 CC -1- FUNCTION: GLANDULAR KALLIKREIN CLEAVE MET-LYS AND ARG-SER BONDS  
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ.  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
 CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
 CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
 CC MET-1-XAA OR LEU-1-XAA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M17962; AAA37541.1; -  
 DR EMBL: M17985; AAA37681.1; -  
 DR EMBL: M17983; AAA37681.1; JOINED.  
 DR EMBL: M17984; AAA37681.1; JOINED.  
 DR EMBL: M18588; AAA39350.1; ALT\_SEQ.  
 DR EMBL: M18608; AAA39351.1; -  
 DR PIR: A29745; A29745.  
 DR PIR: C29746; C29746.  
 DR HSSP: P00757; 1SGF.  
 DR MGD: MGI:95293; EGFBP3.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE: PS00135; TRYPsin\_SER; 1.  
 DR PFAM: PF00089; trypsin; 1.  
 DR Hydroxylase: Serine protease; Kininogenase; Glycoprotein;  
 DR Multigene family: zymogen; Signal.  
 DR FT SIGNAL 1 18  
 DR FT PROPEP 19 24  
 DR FT CHAIN 25 261  
 DR FT ACT\_SITE 65 65  
 DR FT ACT\_SITE 120 120  
 DR FT ACT\_SITE 213 213  
 DR FT DISULFID 31 173  
 DR FT DISULFID 50 66  
 DR FT DISULFID 152 219  
 DR FT DISULFID 198 198  
 DR FT DISULFID 209 234  
 DR FT CARBOHYD 102 102  
 DR SEQUENCE 261 AA; 28900 MW; 796FDADE5620D7 CRC64;  
 SO

Query Match 45.9%; Score 84; DB 1; Length 261;  
 Best Local Similarity 42.9%; Pred. No. 2,59e-05;  
 Matches 9; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Db 165 FONARDLOCVNLKLPNEDCG 185  
 1 : | | | | : | : | : | :  
 Qy 3 FLTPKRLQCVDLHVSNDVCA 23

RESULT 11  
 ID K1KA-RAT STANDARD: PRT: 244 AA.  
 AC P36375;  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE GLANDULAR KALLIKREIN 10 PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN)  
 DE (T-KININOGENASE) (K10) (PROTEINASE B) (ENDOPEPTIDASE K) (FRAGMENT).  
 GN K1K10 OR K1K-10.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE-SUBMANDIBULAR GLAND, AND KIDNEY;  
 RX MEDLINE: 93041794.  
 RA Ma J.-X., Chao J., Chao L.;  
 RT Molecular cloning and characterization of rK1K10, a cDNA encoding T-  
 RT kininogenase from rat submandibular gland and kidney.";  
 RL Biochemistry 31:10922-10928(1992).  
 RN [2]  
 RP SEQUENCE OF 10-32: 95-124 AND 179-232.  
 RA MEDLINE: 91224135.  
 RA Gutman N., Elmoujahed A., Brillard M., du Sorbier B., Gauthier F.;  
 RT "Microheterogeneity of rat submaxillary gland kallikrein K10, a  
 RT member of the kallikrein family.";  
 RL Eur. J. Biochem. 197:425-429(1991).  
 RN [3]  
 RP SEQUENCE OF 10-32 AND 97-133.  
 RC TISSUE-SUBMANDIBULAR GLAND;  
 RX MEDLINE: 90153911.  
 RA Xiong W., Chen L.-M., Chao J.;  
 RT "Purification and characterization of a kallikrein-like  
 RT T-kininogenase.";  
 RL J. Biol. Chem. 265:2822-2827(1990).  
 RN [4]  
 RP SEQUENCE OF 10-32 AND 97-117.  
 RX MEDLINE: 88198057.  
 RA Kato H., Nakaniishi E., Enryoji K., Hayashi I., Oh-Ishi S., Iwanaga S.;  
 RT "Characterization of serine proteinases isolated from rat  
 RT submaxillary gland: with special reference to the degradation of rat  
 RT kininogens by these enzymes.";  
 RL J. Biochem. 102:1389-1404(1987).  
 CC -I- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ. THIS PROTEIN MAY BE  
 CC INVOLVED IN THE REGULATION OF RENAL FUNCTION.  
 CC -I- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
 CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
 CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
 CC MET-1-XAA OR LEU-1-XAA.  
 CC -I- SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND HEAVY CHAIN LINKED  
 CC BY A DISULFIDE BOND.  
 CC -I- TISSUE SPECIFICITY: KIDNEY AND SUBMANDIBULAR GLAND, WHERE IT IS  
 CC FOUND IN THE GRANULAR CONVOLUTED TUBULE AND STRIATED DUCT CELLS.  
 CC IT IS LIKELY THAT THE ENZYME IS MAINLY SYNTHESIZED IN THE GRANULAR  
 CC CONVOLUTED TUBULES AND THEN TRANSFERRED TO OTHER TISSUES BY  
 CC RELEASE INTO THE VASCULATURE OR INTERSTITIAL SPACE.  
 CC -I- PTM: PROBABLY N- AND O-GLYCOSYLATED. IT HAS CARBOHYDRATE MOIETIES  
 CC OF ALPHA-METHYL-D-MANNOSIDE AND N-ACETYL-D-GLUCOSAMINE GROUPS.  
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE  
 CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC -----  
 CC This swiss-prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation.  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: S48142; AAB24071.1; -  
 DR PIR: A35545; A35545.  
 DR PIR: B35545; B35545.  
 DR PIR: A44284; A44284.  
 DR HSSP: P00759; 1TON.  
 DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE: PS00135; TRYPsin\_SER; 1.  
 DR PFAM: PF00089; trypsin; 1.  
 KW Hydroxylase; Serine protease; Kininogenase; Glycoprotein; Signal;  
 KW Multigene family.  
 FT SIGNAL 1 1  
 FT NON\_TER 1 1  
 FT PROPER 4 9 PROBABLE.  
 FT CHAIN 10 244 ACTIVATION PEPTIDE (PROBABLE).  
 FT CHAIN 10 96 T-KININOGENASE LIGHT CHAIN.  
 FT CHAIN 97 244 T-KININOGENASE HEAVY CHAIN.  
 FT ACT\_SITE 48 48 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 103 103 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 196 196 CHARGE RELAY SYSTEM.  
 FT DISULFID 16 156 BY SIMILARITY.  
 FT DISULFID 33 49 BY SIMILARITY.  
 FT DISULFID 135 202 BY SIMILARITY.  
 FT DISULFID 167 181 BY SIMILARITY.  
 FT DISULFID 192 217 BY SIMILARITY.  
 FT CARBOHYD 91 91 POTENTIAL.  
 FT CARBOHYD 76 76 POTENTIAL.  
 FT CARBOHYD 91 91 POTENTIAL.  
 FT CARBOHYD 126 126 POTENTIAL.  
 FT CARBOHYD 142 142 POTENTIAL.  
 FT CONFLICT 29 29 N -> IET (IN REF. 3).  
 FT CONFLICT 115 116 IT -> DS (IN REF. 4).  
 FT CONFLICT 128 128 E -> G (IN REF. 3).  
 FT CONFLICT 133 133 S -> G (IN REF. 3).  
 SO SEQUENCE 244 AA; 27305 MW; BAB4D0547EB79C0 CRC64;

Query Match 45.4%; Score 83; DB 1; Length 244;  
 Best Local Similarity 47.4%; Pred. No. 4,41e-05;  
 Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Db 151 PDDLOCVNLHLSNEKIC 169  
 1 : | | | | : | : | : | :  
 Qy 6 PKRLQCVDLHVSNDVCAQ 24

RESULT 12  
 ID K1K2-RAT STANDARD: PRT: 259 AA.  
 AC P00759;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE TONIN PRECURSOR (EC 3.4.21.35) (ESTERASE 1) (S2 KALLIKREIN) (RGK-2)  
 DE (RSRG-5).  
 GN K1K2 OR K1K-2 OR TON.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 86051477.  
 RA Ashley P.L., McDonald R.J.;  
 RT Kallikrein-related mrnas of the rat submaxillary gland: nucleotide  
 RT sequences of four distinct types including tonin.";  
 RL Biochemistry 24:4512-4520(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89214217.  
 RA Wines D.R., Brady J.M., Pritchett D.B., Roberts J.L., MacDonald R.J.;

RT "Organization and expression of the rat kallikrein gene family."  
 RL J. Biol. Chem. 264:7653-7662(1989).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89375248.  
 RA Shai S.Y., Woodley-Miller C., Chao J., Chao L.  
 RT "Characterization of genes encoding rat contin and a kallikrein-like  
 RL serine protease."  
 RL Biochemistry 28:5334-5343(1989).  
 RN (4)  
 RP SEQUENCE OF 25-259.  
 RX MEDLINE: 87271223.  
 RA Lazure C., Leduc R., Seldah N.G., Thibault G., Genest J., Chretien M.  
 RT "The complete amino acid sequence of rat submaxillary gland tonin  
 RT does contain the aspartic acid at the active site: confirmation by  
 RT protein sequence analysis."  
 RL Biochem. Cell Biol. 65:321-337(1987).  
 RN (5)  
 RP SEQUENCE OF 25-103 AND 120-259.  
 RX MEDLINE: 84117504.  
 RA Lazure C., Leduc R., Seldah N.G., Thibault G., Genest J., Chretien M.  
 RT "Amino acid sequence of rat submaxillary tonin reveals similarities  
 RT to serine proteases."  
 RL Nature 307:555-558(1984).  
 RN (6)  
 RP SEQUENCE OF 25-34.  
 RX MEDLINE: 90147705.  
 RA Kamada M., Furuhata N., Yamaguchi T., Ikekita M., Kizuki K.,  
 RA Moriya H.  
 RT "Observation of tissue prokallikrein activation by some serine  
 RT proteases, arginine esterases in rat submandibular gland."  
 RL Biochem. Biophys. Res. Commun. 166:231-237(1990).  
 RN (7)  
 RP SEQUENCE OF 25-50, AND CHARACTERIZATION.  
 RX MEDLINE: 92250562.  
 RA Moreau T., Brillard-Bourdet M., Bouhnik J., Gauthier F.  
 RT "Protein products of the rat kallikrein gene family. Substrate  
 RT specificities of kallikrein rk2 (tonin) and kallikrein rk9."  
 RL J. Biol. Chem. 267:10045-10051(1992).  
 RN (8)  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
 RX MEDLINE: 88011252.  
 RA Fujinaga M., James M.N.G.  
 RT "Rat submaxillary gland serine protease, tonin. Structure solution  
 RT and refinement at 1.8-A resolution."  
 RL J. Mol. Biol. 195:373-396(1987).  
 CC -1- FUNCTION: THIS PROTEIN HAS BOTH TRYPSIN- AND CHYMOTRYPSIN-LIKE  
 CC ACTIVITIES, BEING ABLE TO RELEASE ANGIOTENSIN II FROM ANGIOTENSIN  
 CC I OR ANGIOTENSINOGEN.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- TISSUE SPECIFICITY: FOUND IN SUBMAXILLARY GLAND.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: M1565; AAA41466.1; -  
 DR EMBL: M23878; AAA42259.1; -  
 DR EMBL: M23877; AAA42259.1; JOINED.  
 DR EMBL: M26533; AAA42081.1; -  
 DR PIR: A00945; KORTIN.  
 DR PIR: A30971; A30971.  
 DR PIR: A34050; A34050.  
 DR PIR: C23863; C23863.  
 DR PIR: B33359; B33359.  
 DR PIR: A32340; A32340.  
 DR PDB: 1TON; 16-JAN-88.

DR PRINTS: P00722; CHYMOTRYPSIN.  
 DR PROSITE: P500134; TRYPSIN\_HIS; 1.  
 DR PROSITE: P500135; TRYPSIN\_SER; 1.  
 DR PFAM: P000089; trypsin; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Multigene family; Zymogen;  
 FT Signal; 3d-structure.  
 FT SIGNAL 1  
 FT PROPEP 19 24  
 FT CHAIN 25 259  
 FT ACT\_SITE 63 63  
 FT ACT\_SITE 118 118  
 FT ACT\_SITE 211 211  
 FT ACT\_SITE 31 171  
 FT DISULFID 48 64  
 FT DISULFID 150 217  
 FT DISULFID 182 196  
 FT DISULFID 207 232  
 FT CARBOHYD 106 106  
 FT CARBOHYD 189 189  
 FT STRAND 26 26  
 FT STRAND 29 30  
 FT TURN 33 34  
 FT TURN 37 38  
 FT STRAND 39 43  
 FT STRAND 47 54  
 FT STRAND 55 56  
 FT STRAND 57 60  
 FT HELIX 62 64  
 FT STRAND 70 73  
 FT STRAND 77 77  
 FT TURN 78 79  
 FT TURN 83 84  
 FT STRAND 86 88  
 FT STRAND 90 95  
 FT TURN 97 98  
 FT TURN 116 117  
 FT STRAND 120 124  
 FT TURN 146 147  
 FT STRAND 149 154  
 FT STRAND 168 168  
 FT STRAND 170 177  
 FT HELIX 179 181  
 FT HELIX 183 186  
 FT TURN 188 189  
 FT HELIX 190 193  
 FT STRAND 194 198  
 FT TURN 200 201  
 FT STRAND 205 205  
 FT TURN 208 209  
 FT TURN 211 212  
 FT STRAND 214 217  
 FT TURN 218 219  
 FT STRAND 220 225  
 FT TURN 235 236  
 FT STRAND 239 243  
 FT HELIX 244 257  
 FT SEQUENCE 259 AA; 28248 MW; 3D6E6D011F926B4 CRC64;  
 SO Query Match 44.88; Score 82; DB 1; Length 259;  
 Best Local Similarity 39.18; Pred. No. 7,49e-05;  
 Matches 9; Conservative 8; Mismatches 6; Indels 0; Gaps 0;  
 Db 162 EMBVSHDLOCVNIHLISNEKCE 184  
 Oy 2 EFLTPKRLQCVDLHVLSNDVCAO 24  
 RESULT 13  
 ID KIK3\_RAT STANDARD: PRT: 188 AA.  
 AC P15950;  
 DT 01-APR-1990 (rel. 14, Created)  
 DT 01-JUN-1994 (rel. 29, Last sequence update)  
 DT 01-NOV-1997 (rel. 35, Last annotation update)

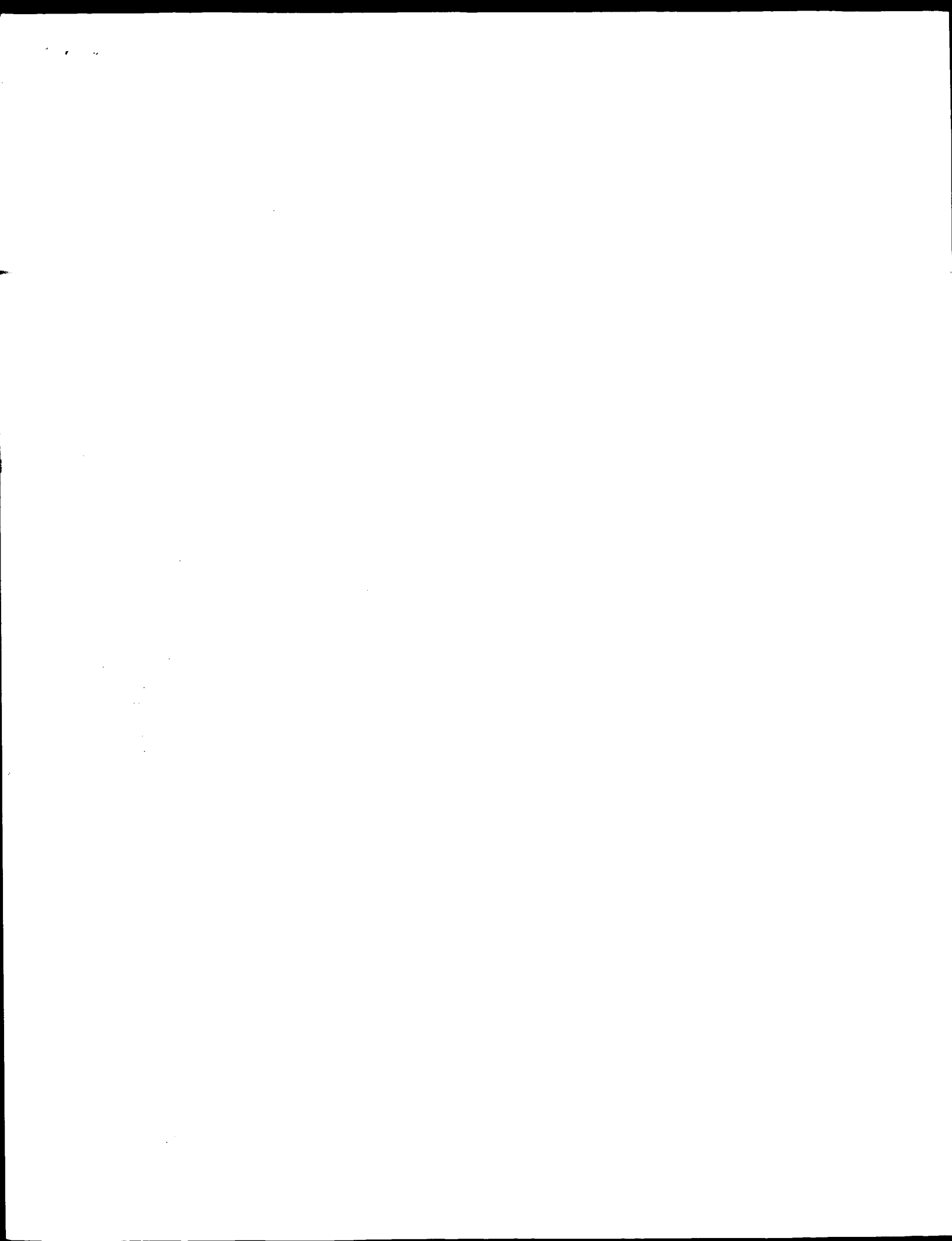
DE GLANDULAR KALLIKREIN 3, SUBMANDIBULAR (EC 3.4.21.35) (TISSUE  
DE KALLIKREIN) (S1 KALLIKREIN) (RKG-3) (RSKG-50) (FRAGMENT).  
GN KLK3 OR KLK-3.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89375248.  
RA Shai S.Y., Woodley-Miller C., Chao J., Chao L.;  
RT "Characterization of genes encoding rat tonin and a kallikrein-like  
RT serine protease."  
RL Biochemistry 28:5334-5343(1989).  
RN [2]  
RP SEQUENCE OF 33-188 FROM N.A.  
RX MEDLINE; 86051477.  
RA Ashley P.L., Macdonald R.J.;  
RT "Kallikrein-related mrnas of the rat submaxillary gland: nucleotide  
RT sequences of four distinct types including tonin."  
RL Biochemistry 24:4512-4520(1985).  
CC -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
CC IN KININOGEN TO RELEASE LYS-BRADYKININ.  
CC CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
CC MET-1-XAA OR LEU-1-XAA.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M26534; AAA42080.1; -;  
DR EMBL; M15564; AAA41465.1; -;  
DR PIR; B23863; B23863.  
DR PIR; B32340; B32340.  
DR HSSP; P00759; 1TON.  
DR PROSITE; PS00134; TRYPsin\_HIS; PARTIAL.  
DR PROSITE; PS00135; TRYPsin\_SER; 1.  
DR PRAM; PF00089; trypsin.1.  
KW Hydrolase; Serine protease; Kininogenase; Glycoprotein;  
KM Multigene family.  
FT NON\_TER 1  
FT ACT\_SITE 47 1 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 140 140 CHARGE RELAY SYSTEM.  
FT DISULFID 79 146 BY SIMILARITY.  
FT DISULFID 111 125 BY SIMILARITY.  
FT DISULFID 136 161 BY SIMILARITY.  
FT DISULFID 34 35 MR -> IW (IN REF. 2).  
FT CONFLICT 42 42 K -> G (IN REF. 2).  
FT CONFLICT 186 186 E -> K (IN REF. 1).  
SQ SEQUENCE 188 AA; 20986 MW; B3CECE65582EB1E9 CRC64;  
Query Match 44.3%; Score 81; DB 1; Length 188;  
Best Local Similarity 52.9%; Pred. No. 1.27e-04;  
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

DE GLANDULAR KALLIKREIN 8, PROSTATIC PRECURSOR (EC 3.4.21.35) (TISSUE  
DE KALLIKREIN) (P1 KALLIKREIN) (RKG-8).  
GN KLK8 OR KLK-8.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89352606.  
RA Brady J.M., Wines D.R., Macdonald R.J.;  
RT "Expression of two kallikrein gene family members in the rat  
RT prostate."  
RL Biochemistry 28:5203-5210(1989).  
CC -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS  
CC IN KININOGEN TO RELEASE LYS-BRADYKININ.  
CC CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
CC MET-1-XAA OR LEU-1-XAA.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPsin FAMILY. GLANDULAR KALLIKREIN SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M27217; AAA42036.1; -;  
DR EMBL; M27215; AAA42036.1; JOINED.  
DR EMBL; M27216; AAA42036.1; JOINED.  
DR PIR; A34079; A34079.  
DR HSSP; P00759; 1TON.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
DR PROSITE; PS00135; TRYPsin\_SER; 1.  
DR PRAM; PF00089; trypsin.1.  
KW Hydrolase; Serine protease; Kininogenase; Glycoprotein;  
KM Multigene family; Zymogen; Signal.  
FT SIGNAL 1 18 PROBABLE.  
FT PROPEP 19 24 ACTIVATION PEPTIDE (PROBABLE).  
FT CHAIN 25 261 GLANDULAR KALLIKREIN 8, PROSTATIC.  
FT ACT\_SITE 65 65 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM.  
FT DISULFID 31 173 BY SIMILARITY.  
FT DISULFID 50 66 BY SIMILARITY.  
FT DISULFID 152 219 BY SIMILARITY.  
FT DISULFID 184 198 BY SIMILARITY.  
FT DISULFID 209 234 BY SIMILARITY.  
FT CARBOHYD 108 108 POTENTIAL.  
SQ SEQUENCE 261 AA; 29013 MW; CA8F12151B04E337 CRC64;  
Query Match 44.3%; Score 81; DB 1; Length 261;  
Best Local Similarity 52.9%; Pred. No. 1.27e-04;  
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

RESULT 14  
ID KLK8\_RAT STANDARD; PRT; 261 AA.  
AC P36374;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)

RESULT 15  
ID KLK7\_RAT STANDARD; PRT; 261 AA.  
AC P36373;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE GLANDULAR KALLIKREIN 7, SUBMANDIBULAR/RENAL PRECURSOR (EC 3.4.21.35)  
DE (TISSUE KALLIKREIN) (RKG-7) (RSKG-7) (ESTERASE B) (PROTEINASE A).  
GN KLK7 OR KLK-7.







```

OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92324494.
RA RIEGMAN P.H., VILLETSTRA R.J., VAN DER KORPUT H.A., ROMIJN J.C.,
RT "Identification and androgen-regulated expression of two major human
RT glandular kallikrein-1 (hK-1) mRNA species."
RL Mol. Cell. Endocrinol. 76:181-190(1991).
DR EMBL: S39329; AADI3817.1; -.
DR HSSP: P07288; 1PFA.
DR PFAM: PF00089; trypsin; 1.
SQ SEQUENCE 223 AA; 24674 MW; E0321503 CRC32;

Query Match
Best Local Similarity 66.7%; Score 130; DB 4; Length 223;
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 163 EEF1PKKLCQVDLHVISNDVCA 186
QY 1 EEF1PKKLCQVDLHVISNDVCA 24

RESULT 3
ID 029474 PRELIMINARY: PRT; 261 AA.
AC 029474;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE KALLIKREIN PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN)
DE (GLANDULAR KALLIKREIN) (KININOGENIN).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-MONGREL;
RX MEDLINE: 94250683.
RA GAUTHIER E.R., DUMAS C., CHAPDELAINE P., TREMBLAY R.R., DUBE J.Y.;
RT "Characterization of canine pancreas kallikrein cDNA."
RL Biochim. Biophys. Acta 1218:102-104(1994).
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE
CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF
CC MET-1-XAA OR LEU-1-XAA. THE RAT ENZYME IS UNUSUAL IN LIBERATING
CC BRADYKININ DIRECTLY FROM AUTOLOGOUS KININOGENS BY CLEAVAGE AT TWO
CC ARG-1-XAA BONDS.
DR EMBL: X75479; CA53210.1; -.
DR HSSP: P00752; 2KAI.
DR PFAM: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
KW Signal; Hydrolase.
FT SIGNAL 1
FT CHAIN 25 24 POTENTIAL.
FT NON_TER 52 52 KALLIKREIN.
SQ SEQUENCE 261 AA; 28913 MW; 65B49AEF CRC32;

Query Match
Best Local Similarity 59.0%; Score 108; DB 6; Length 261;
Matches 13; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 163 DKFIIPDLQCVDELISNDICA 185
QY 1 EEF1PKKLCQVDLHVISNDVCA 23

RESULT 4
ID 016272 PRELIMINARY: PRT; 189 AA.
AC 016272;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE PSA PROTEIN (FRAGMENT).
OS Eukaryota; Primates; Catarrhini; Homiidae; Homo.

```

---

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95079406.
RA MONNE M., CROCE C.M., YU H., DIAMANDIS E.P.;
RT "Molecular characterization of prostate-specific antigen messenger RNA
RT expressed in breast tumors."
RL Cancer Res. 54:6344-6347(1994).
DR EMBL: S75755; AADI4185.1; -.
DR HSSP: P07288; 1PFA.
DR PFAM: PF00089; trypsin; 1.
FT NON_TER 1
SQ SEQUENCE 189 AA; 21039 MW; DBAC6482 CRC32;

Query Match
Best Local Similarity 100.0%; Score 107; DB 4; Length 189;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 176 EEF1PKKLCQVDL 189
QY 1 EEF1PKKLCQVDL 14

RESULT 5
ID 046630 PRELIMINARY: PRT; 52 AA.
AC 046630;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE TISSUE KALLIKREIN (EC 3.4.21.35) (GLANDULAR KALLIKREIN) (KININOGENIN)
DE (FRAGMENT)
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
OC Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RX DEDIO J., WOHLEFART P., WIEMER G., SCHOLKENS B.A., MULLER-ESTERL W.;
RT Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
RL -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN
CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE
CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF
CC MET-1-XAA OR LEU-1-XAA. THE RAT ENZYME IS UNUSUAL IN LIBERATING
CC BRADYKININ DIRECTLY FROM AUTOLOGOUS KININOGENS BY CLEAVAGE AT TWO
CC ARG-1-XAA BONDS.
DR EMBL: AJ003191; CA05970.1; -.
DR HSSP: P00752; 2KAI.
KW Hydrolase.
FT SIGNAL 1
FT CHAIN 52 52
FT NON_TER 52 52
SQ SEQUENCE 52 AA; 5763 MW; C9EB33B7 CRC32;

Query Match
Best Local Similarity 52.2%; Score 92; DB 6; Length 52;
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db 5 DEFSPYDLCQVDLILPNEKCA 27
QY 1 EEF1PKKLCQVDLHVISNDVCA 23

RESULT 6
ID 007277 PRELIMINARY: PRT; 195 AA.
AC 007277;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE KALLIKREIN PRECURSOR (EC 3.4.21.35) (TISSUE KALLIKREIN).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

```



RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LIVER:  
 RA MEDLINE: 93305727.  
 RA LIN F.R., LIN C.H., CHOU C., CHEN K., LU H.S., BACHELLER B.,  
 RA HERRERA C., JONES T., CHAO J., CHAO L.,  
 RT "Molecular cloning and sequence analysis of the monkey and human  
 RT tissue kallikrein genes."  
 RL Blochim. Biophys. Acta 1173:325-328(1993).  
 DR EMBL: L10038; AAA36135.1; -.  
 DR HSSP: P07288; 1PPA.  
 DR PFAM: PF00089; trypsin. 1.  
 KW Hydrolyase; Serine protease; Kininogenase; Glycoprotein; Zymogen;  
 KN Signal.  
 FT PROPEP 1 2 ACTIVATION PEPTIDE.  
 FT SIGNAL 1 2  
 FT CHAIN 1 2  
 FT ACT\_SITE 147 147 KALLIKREIN.  
 FT DISULFID 86 153 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 118 132 BY SIMILARITY.  
 FT DISULFID 143 168 BY SIMILARITY.  
 FT CARBOHYD 98 98 POTENTIAL.  
 FT CARBOHYD 100 100 POTENTIAL.  
 SQ SEQUENCE 195 AA; 21091 MW; 0DIE38EC CRC32;  
 Query Match 49.7%; Score 91; DB 4; Length 195;  
 Best Local Similarity 54.5%; Pred. No. 5.98e-07;  
 Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 Db 97 ENFSPDLCVLDKILPNDK 118  
 1 EELTPKLLQCVLDHVSNDVC 22  
 OY 1 EELTPKLLQCVLDHVSNDVC 22  
 RESULT 7  
 ID 003955 PRELIMINARY; PRT; 250 AA.  
 AC 003955;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBLrel. 12, Last annotation update)  
 DE GLANDULAR KALLIKREIN (FRAGMENT).  
 OS Praxmys natalensis (African soft-furred rat) (Mastomys natalensis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mastomys.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SALIVARY GLAND;  
 RX MEDLINE: 94226702.  
 RA BELL R.A., FAHNESTOCK M.,  
 RT "Characterization of kallikrein cDNAs from the African rodent  
 RT Mastomys."  
 RL DNA Cell Biol. 13:293-300(1994).  
 CC -1- FUNCTION: GLANDULAR KALLIKREIN CLEAVE MET-LYS AND ARG-SER BONDS  
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ.  
 DR EMBL: X17351; CAA35231.1; -.  
 DR HSSP: P36368; 1AOS.  
 DR PFAM: PF00089; trypsin. 1.  
 KW Hydrolyase; Serine protease; Kininogenase; Glycoprotein; Saliva;  
 KW Submandibular gland; Multigene family; Zymogen; Signal.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 250 AA; 27385 MW; BD8435E0 CRC32;  
 Query Match 47.5%; Score 87; DB 11; Length 250;  
 Best Local Similarity 47.8%; Pred. No. 5.54e-06;  
 Matches 11; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
 Db 161 EEYKFPDLCVYKLPNFEICA 183  
 1 EELTPKLLQCVLDHVSNDVC 23  
 OY 1 EELTPKLLQCVLDHVSNDVC 23  
 RESULT 8  
 ID 09Y5K2 PRELIMINARY; PRT; 254 AA.

AC 09Y5K2;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE KALLIKREIN 4.  
 GN KLR4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STEPHENSON S.A., VERTY K., ASHWORTH L., CLEMENTS J.A.,  
 RT "Localization of a new prostate specific antigen-related serine  
 RT protease gene, KLR4, is evidence for an expanded human kallikrein  
 RT (KLK) gene family cluster on chromosome 19q13.3-13.4."  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF148532; AAD38019.1; -.  
 SQ SEQUENCE 254 AA; 27022 MW; 0CFAD39C CRC32;  
 Query Match 47.0%; Score 86; DB 4; Length 254;  
 Best Local Similarity 45.0%; Pred. No. 9.60e-06;  
 Matches 9; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
 Db 162 PTVLCVNVSVSEVCSKL 181  
 6 PKRLQCVLDHVSNDVCAY 25  
 OY 6 PKRLQCVLDHVSNDVCAY 25  
 RESULT 9  
 ID 063275 PRELIMINARY; PRT; 239 AA.  
 AC 063275;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBLrel. 12, Last annotation update)  
 DE KALLIKREIN (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-SUBMANDIBULAR GLAND;  
 RA ZINTZ C.B., MA J.X., CHAO J., CHAO L.,  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: L33840; AAA58782.1; -.  
 DR HSSP: P00759; 1TON.  
 DR PFAM: PF00089; trypsin. 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 239 AA; 26382 MW; 0609E3B6 CRC32;  
 Query Match 44.3%; Score 81; DB 11; Length 239;  
 Best Local Similarity 52.9%; Pred. No. 1.44e-04;  
 Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 Db 146 PDDLCVNIHLNNEKC 162  
 6 PKRLQCVLDHVSNDVC 22  
 OY 6 PKRLQCVLDHVSNDVC 22  
 RESULT 10  
 ID 016101 PRELIMINARY; PRT; 265 AA.  
 AC 016101;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE SERINE PROTEASE SER4 PRECURSOR.  
 GN SER4.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OREGON R;

RX MEDLINE: 98146263.  
 RA ARENS J.E., MAHONEY P.A.:  
 RT "Isolation and cloning of Ser4, a gene encoding a trypsin-like serine  
 RT protease in Drosophila melanogaster."  
 RL Biochim. Biophys. Acta 1395:141-144(1998).  
 DR EMBL: AF006639; AAC14351.1; -  
 DR HSSP: P08897; 1HYL.  
 DR FLYBASE: FBgn0020906; Ser4.  
 DR PFM: PFM0089; trypsin.1.  
 KW Signal; Protease.  
 FT SIGNAL 1 21  
 FT CHAIN 37 265  
 SQ SEQUENCE 265 AA; 28950 MW; 5272F05B CRC32;  
 Query Match 43.7%; Score 80; DB 5; Length 265;  
 Best Local Similarity 47.4%; Pred. No. 2,45e-04;  
 Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
 Db 171 PDMMECVLDIISNCSR 189  
 QY 6 PKKLCVLDLHVISNDVC 24  
 RESULT 11  
 ID 09XSN6 PRELIMINARY; PRT; 254 AA.  
 AC 09XSN6;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE ENAMEL MATRIX SERINE PROTEINASE 1 PRECURSOR.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA SIMMER J.P., FUKAE M., TANABE T., YAMAKOSHI Y., UCHIDA T., XUE J.,  
 RA MARGOLIS H.C., SHIMIZU M., HU C.-C., BARTLETT J.D.:  
 RT "Purification, Characterization and Cloning of Enamel Matrix Serine  
 RT Proteinase 1."  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U76256; AAB94638.1; -  
 KW Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 31 254  
 SQ SEQUENCE 254 AA; 27235 MW; 256DAA2 CRC32;  
 Query Match 42.6%; Score 78; DB 6; Length 254;  
 Best Local Similarity 52.9%; Pred. No. 7.02e-04;  
 Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 Db 162 PVLQCVNISVASEVC 178  
 QY 6 PKKLCVLDLHVISNDVC 22  
 RESULT 12  
 ID 061855 PRELIMINARY; PRT; 261 AA.  
 AC 061855;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE TISSUE KALLIKREIN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TADA M., PETERS J., TAKAHASHI S., INOUE H., MIYAKE Y.:  
 RT "Identification of a tissue kallikrein gene, mck-6, expressed in a  
 RT mouse neuroendocrine cell line."  
 RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: D10464; BAA01257.1; -

DR HSSP: P36368; 1AOS.  
 DR PFM: PFM0089; trypsin.1.  
 DR PRINTS: PRO0722; CHYMOTRYPSIN.  
 SQ SEQUENCE 261 AA; 28775 MW; 509D1D00 CRC32;  
 Query Match 42.6%; Score 78; DB 11; Length 261;  
 Best Local Similarity 50.0%; Pred. No. 7.02e-04;  
 Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
 Db 168 PDELQCVNLKLPNEDCA 185  
 QY 6 PKKLCVLDLHVISNDVC 23  
 RESULT 13  
 ID 061761 PRELIMINARY; PRT; 96 AA.  
 AC 061761;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE KALLIKREIN (FRAGMENT).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MASON A.D., EVANS B.A., COX D.R., SHINE J., RICHARDS R.I.:  
 RT "Structure of mouse kallikrein gene family suggests a role in specific  
 RT processing of biologically active peptides."  
 RL Nature 303:300-307(1983)  
 DR EMBL: V00829; CAA24212.1; -  
 DR HSSP: P36368; 1AOS.  
 FT NON\_TER 1  
 SQ SEQUENCE 96 AA; 10327 MW; CDDC7116 CRC32;  
 Query Match 41.5%; Score 76; DB 11; Length 96;  
 Best Local Similarity 47.1%; Pred. No. 1.99e-03;  
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
 Db 3 AKDLQCVNLKLPNENC 19  
 QY 6 PKKLCVLDLHVISNDVC 22  
 RESULT 14  
 ID 0920M1 PRELIMINARY; PRT; 255 AA.  
 AC 0920M1;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DE ENAMEL MATRIX SERINE PROTEINASE 1 PRECURSOR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-SWISS-WEBSTER;  
 RA SIMMER J.:  
 RT "Enamel Matrix Serine Proteinase 1 (EMSP1)."  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF019979; AAC98894.1; -  
 DR HSSP: P00763; 1DPO.  
 KW Signal.  
 FT SIGNAL 1 25  
 FT CHAIN 32 255  
 SQ SEQUENCE 255 AA; 27488 MW; 6E711616 CRC32;  
 Query Match 41.0%; Score 75; DB 11; Length 255;  
 Best Local Similarity 52.9%; Pred. No. 3.33e-03;  
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Db 163 PSLQCVNLISVASEVC 179

OY 6 PKKLOCVDLHVISNDVC 22

RESULT 15  
ID 063274 PRELIMINARY; PRT; 235 AA.

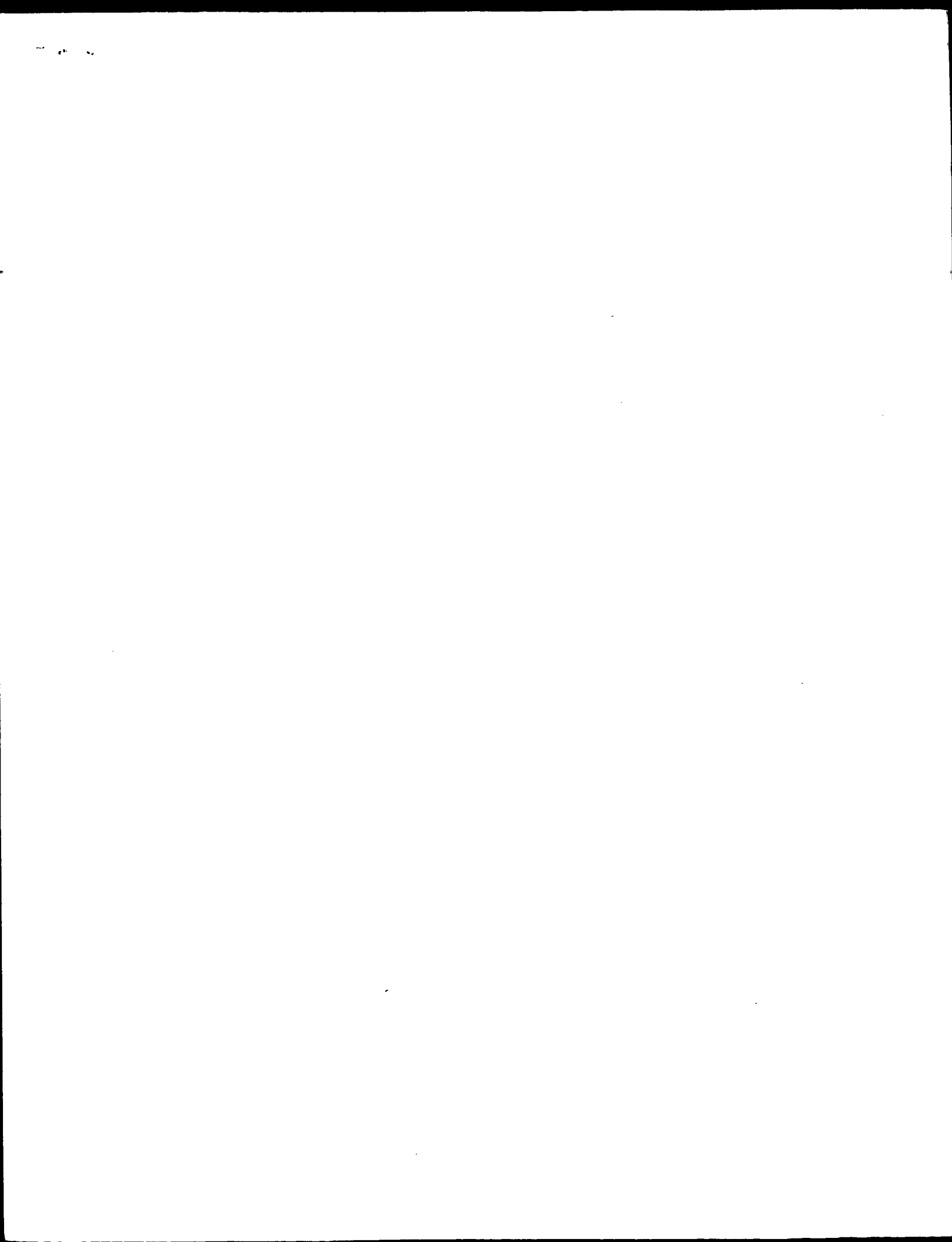
AC 063274;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
DE KALIKREIN (FRAGMENT).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

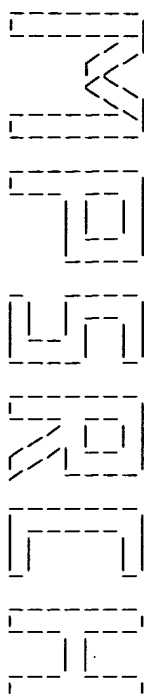
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-KIDNEY;  
RA 2INTZ C.B., MA J.X., CHAO J., CHAO L.;  
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L33839; AAA58781.1; -.  
DR HSSP; P00759; ITON.  
DR PFAM; PF00089; trypsin; 1.  
FT NON\_TER  
SO SEQUENCE 235 AA; 26226 MM; E6DE5AD8 CRC32;

Query Match 40.4%; Score 74; DB 11; Length 235;  
Best Local Similarity 42.1%; Pred. No. 5.55e-03;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

DB 142 PDDLCCVNDLSENKIC 160  
OY 6 PKKLOCVDLHVISNDVC 24

Search completed: Thu Jun 22 21:22:40 2000  
Job time : 14 secs.





(TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsearch n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 15:53:38 2000; Maspar time 4470.62 seconds  
67.998 Million cell updates/sec

Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTACNCCNNA.....AYGATGNTGTCNCARGTN 75  
Comp: CTCTCTAARRANTGNGNTT.....TCTCTACNACRCNGTTCAN

Scoring table:  
TABLE bkttranslated2  
Gap 40

Mismatch STD : Dbase 0; Query 0

Searched: 4857316 seqs, 2026611650 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database:

embl-est61  
1:em\_est11 2:em\_est12 3:em\_est13 4:em\_est14 5:em\_est15  
6:em\_est16 7:em\_est17 8:em\_est18 9:em\_est19 10:em\_est20  
11:em\_est21 12:em\_est22 13:em\_est23 14:em\_est24 15:em\_est25  
16:em\_gss12 17:em\_gss2 18:em\_gss3 19:em\_gss4 20:em\_gss5  
21:em\_gss6  
genbank-est116  
22:gb\_est1 23:gb\_est10 24:gb\_est11 25:gb\_est12  
26:gb\_est13 27:gb\_est14 28:gb\_est15 29:gb\_est16  
30:gb\_est17 31:gb\_est18 32:gb\_est19 33:gb\_est20  
34:gb\_est21 35:gb\_est22 36:gb\_est23 37:gb\_est24  
38:gb\_est25 39:gb\_est26 40:gb\_est27 41:gb\_est28  
42:gb\_est29 43:gb\_est30 44:gb\_est31 45:gb\_est32  
46:gb\_est33 47:gb\_est34 48:gb\_est35 49:gb\_est36  
50:gb\_est37 51:gb\_est38 52:gb\_est39 53:gb\_est40  
54:gb\_est41 55:gb\_est42 56:gb\_est43 57:gb\_est44  
58:gb\_est45 59:gb\_est46 60:gb\_est47 61:gb\_est48  
62:gb\_est49 63:gb\_est50 64:gb\_est51 65:gb\_est52  
66:gb\_est53 67:gb\_est54 68:gb\_est55 69:gb\_est56  
70:gb\_est57 71:gb\_est58 72:gb\_est59 73:gb\_est60  
74:gb\_est61 75:gb\_est62 76:gb\_est63 77:gb\_est64  
78:gb\_est65 79:gb\_est66 80:gb\_est67 81:gb\_est68  
82:gb\_est69 83:gb\_est70 84:gb\_est71 85:gb\_est72  
86:gb\_est73 87:gb\_est74 88:gb\_est75 89:gb\_est76  
90:gb\_est77 91:gb\_est78 92:gb\_est79 93:gb\_est80  
94:gb\_est81 95:gb\_est82 96:gb\_est83 97:gb\_est84  
98:gb\_est85 99:gb\_est86 100:gb\_est87

1	312	98.7	373	24	AA228953	nc14604.r1	NCI CGAP Pr	1.08e-49
2	312	98.7	388	24	AA228953	nc15604.r1	NCI CGAP Pr	1.08e-49
3	312	98.7	420	24	AA228953	nc16006.r1	NCI CGAP Pr	1.08e-49
4	312	98.7	426	24	AA228953	nc17009.r1	NCI CGAP Pr	1.08e-49
5	312	98.7	436	30	AA650872	nc15805.s1	NCI CGAP Pr	1.08e-49
6	312	98.7	439	29	AA650872	nc123611.s1	NCI CGAP Pr	1.08e-49
7	312	98.7	453	29	AA573727	nc13609.s1	NCI CGAP Pr	1.08e-49
8	312	98.7	465	30	AA654527	nc159812.s1	NCI CGAP Pr	1.08e-49
9	312	98.7	471	30	AA654924	nc17607.s1	NCI CGAP Pr	1.08e-49
10	312	98.7	481	29	AA579159	nc12808.s1	NCI CGAP Pr	1.08e-49
11	312	98.7	489	30	AA658261	nc121003.s1	NCI CGAP Pr	1.08e-49
12	312	98.7	491	24	AA225115	nc121011.s1	NCI CGAP Pr	1.08e-49
13	312	98.7	505	29	AA574033	nc146111.s1	NCI CGAP Pr	1.08e-49
14	312	98.7	520	28	AA654295	nc170605.s1	NCI CGAP Pr	1.08e-49
15	312	98.7	542	30	AA654558	nc159011.s1	NCI CGAP Pr	1.08e-49
16	312	98.7	554	29	AA579039	nc134017.s1	NCI CGAP Pr	1.08e-49
17	312	98.7	681	47	AI732097	nc121011.x5	NCI CGAP Pr	1.08e-49
18	312	98.7	732	43	AI547309	PM001_AH.B02.r	Yodnorm	1.08e-49
19	312	98.7	779	43	AI547084	PM02.1.1.G05.r	mynorm	1.08e-49
20	307	97.2	772	22	T29518	EST82635	Human Prostat	2.53e-48
21	302	95.6	398	30	AA653287	nc171007.s1	NCI CGAP Pr	5.87e-47
22	302	95.6	457	51	AI926979	nc168011.x1	NCI CGAP Pr	5.87e-47
23	302	95.6	558	30	AA640352	nc121012.s1	NCI CGAP Pr	5.87e-47
24	292	92.4	487	24	AA229599	nc135002.s1	NCI CGAP Pr	3.05e-44
25	277	87.7	328	30	AA654296	nc183006.s1	NCI CGAP Pr	3.32e-40
26	272	86.1	406	27	AA420570	nc161006.r1	NCI CGAP Pr	7.18e-39
27	272	86.1	500	27	AA420845	nc160004.s1	NCI CGAP Pr	7.18e-39
28	257	81.3	582	29	AA594946	nc140001.s1	NCI CGAP Pr	6.69e-35
29	252	79.7	414	27	AA420522	nc161006.s1	NCI CGAP Pr	1.37e-33
30	242	76.6	329	28	AA522842	nc172008.s1	NCI CGAP Pr	5.45e-31
31	242	76.6	929	29	AA558309	nc135004.s1	NCI CGAP HS	5.45e-31
32	242	76.6	902	42	AI525832	nc130604.s1	NCI CGAP HS	5.45e-31
33	232	73.4	273	30	AA659308	nc179604.s1	NCI CGAP Pr	2.04e-28
34	232	73.4	283	28	AA553453	nc176001.s1	NCI CGAP Pr	2.04e-28
35	232	73.4	377	24	AA226360	nc171009.s1	NCI CGAP Pr	2.04e-28
36	232	73.4	804	43	AA557338	nc171009.s1	NCI CGAP Pr	2.04e-28
37	210	66.5	189	30	AA557663	nc182006.s1	NCI CGAP Pr	7.12e-23
38	191	60.4	104	30	AA657697	nc182006.s1	NCI CGAP Pr	3.11e-18
39	191	60.4	485	29	AA573575	nc141110.s1	NCI CGAP Pr	3.11e-18
40	181	57.3	413	47	AI732674	nc167005.x5	NCI CGAP Pr	7.42e-16
41	181	57.3	428	28	AA469129	nc167005.s1	NCI CGAP Pr	7.42e-16
42	181	57.3	433	29	AA551415	nc155602.s1	NCI CGAP Pr	7.42e-16
43	181	57.3	513	28	AA534246	nc170611.s1	NCI CGAP Pr	7.42e-16
44	181	57.3	537	29	AA55047	nc140103.s1	NCI CGAP Pr	7.42e-16
45	181	57.3	566	30	AA653017	nc189003.s1	NCI CGAP Pr	7.42e-16

## ALIGNMENTS

RESULT 1  
LOCUS AA228953 373 bp mRNA  
DEFINITION nc14604.r1 NCI CGAP Pr1 Homo sapiens cDNA clone IMAGE:1008078  
similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN).  
ACCESSION AA228953  
VERSION AA228953.1 GI:1651772  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 373)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394945.  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuangui,  
M.D., Michael Emmert-Buck, M.D., Ph.D.

Statistics: Mean 53.419; Variance 70.507; scale 0.758  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No. Query Score Match Length DB ID Description Pred. No.



FEATURES  
source

High quality sequence stop: 377.

Location/Qualifiers

1. 420

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_1lb="NCI\_CGAP\_Pr1"

/sex="Male"

/dev\_stage="45 years old"

/lab\_host="DH10B"

/note="Vector: PAMF10; Site\_1: NotI; Site\_2: EcoRI; 1st

strand cDNA was primed with oligo(dT)17 on 50 ng of

DNase-treated, total cellular RNA obtained from

5,000-10,000 microdissected, histologically normal

prostate epithelial cells. Double-stranded cDNA was

ligated to EcoRI adaptors, 5 cycles of PCR applied to the

cDNA with an adaptor-specific primer, and the resulting

PCR product subcloned into PAMF10 by the UDG-cloning

method (Life Technologies). Average insert size is 600

bp. NOTE: Not directionally cloned. This library was

constructed by David Krizman."

BASE COUNT  
92 a 120 c 119 g 89 t

ORIGIN

Query Match 98.7%; Score 312; DB 24; Length 420;

Best Local Similarity 60.8%; Pred. No. 1.08e-49;

Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 97 GAGGAGTCTTGACCCCAAGAACTTCAGTGTGGACCTCCATGTTATTCGAATGAC 156

QY 1 GARGARTTYTNACNCNARARAYTNCARTGTGAYTNCAYGTATHTMSNAAYGAY 60

Db 157 GTGTGTGGCAAGT 170

QY 61 GTTGTGCGCARGT 74

RESULT 4 AA226359 426 bp mRNA EST 20-AUG-1997

LOCUS ncl12h09.r1 NCI\_CGAP\_Pr1 Homo sapiens cDNA clone IMAGE:1008449

DEFINITION similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);

Accession AA226359

VERSION AA226359.1 GI:1847666

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

1 (bases 1 to 426)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

Journal National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished (1997)

On Sep 12, 1996 this sequence version replaced gi:1290383.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,

M.D., Michael Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: David B. Krizman, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/dbxrp/image/image.html

Seq primer: -28m3 rev1 ET from Amersham

High quality sequence stop: 379.

Location/Qualifiers

1. 426

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_1lb="NCI\_CGAP\_Pr1"

/sex="Male"

/dev\_stage="45 years old"

/lab\_host="DH10B"

/note="Vector: PAMF10; Site\_1: NotI; Site\_2: EcoRI; 1st

BASE COUNT  
92 a 110 c 120 g 101 t 3 others

ORIGIN

Query Match 98.7%; Score 312; DB 24; Length 426;

Best Local Similarity 60.8%; Pred. No. 1.08e-49;

Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 78 GAGGAGTCTTGACCCCAAGAACTTCAGTGTGGACCTCCATGTTATTCGAATGAC 137

QY 1 GARGARTTYTNACNCNARARAYTNCARTGTGAYTNCAYGTATHTMSNAAYGAY 60

Db 138 GTGTGTGGCAAGT 151

QY 61 GTTGTGCGCARGT 74

RESULT 5 AA630877 436 bp mRNA EST 20-OCT-1997

LOCUS nt58a05.s1 NCI\_CGAP\_Pr3 Homo sapiens cDNA clone IMAGE:1202672

DEFINITION similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);

Accession AA630877

VERSION AA630877.1 GI:2553488

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

1 (bases 1 to 436)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

Journal National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished (1997)

On Sep 12, 1996 this sequence version replaced gi:1397616.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,

M.D., Michael Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: David B. Krizman, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/dbxrp/image/image.html

Seq primer: -40m3 fwd, ET from Amersham.

High quality sequence stop: 379.

Location/Qualifiers

1. 436

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_1lb="NCI\_CGAP\_Pr3"

/sex="Male"

/dev\_stage="45 years old"

/lab\_host="DH10B"

/note="Vector: PAMF10; Site\_1: NotI; Site\_2: EcoRI; 1st







[illegible]

Query Match	98.7%	Score 312	DB 29	Length 441
Best Local Similarity	60.8%	Pred. No. 1,08e+49		
Matches	45; Conservative	19; Mismatches	10; Indels	0; Gaps
Db	171	GAGGATTTTTCACCCCAAGAACTTCACAGGTGTGACCCATCTTATTTTCCCAATAC	230	
Oy	1	GAGGATTTTTCACCCCAAGAACTTCACAGGTGTGACCCATCTTATTTTCCCAATAC	230	

Db	231	GTGTGGCGCAGT	244
Qy	61	GTGTGGCGCAGT	74
RESULT	11		
LOCUS	AA658261	489 bp	MRNA
DEFINITION	nu12c03.81 NCI-CGAP Pr2 Homo sapiens cDNA clone IMAGE:1208644 similar to gb:u21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);		EST
ACCESSION	AA658261		05-NOV-1997
VERSION	AA658261.1	GI:2594415	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 489)		
TITLE	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .		
	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	On May 8, 1995 this sequence version replaced gi:801261. Contact: Robert Strausberg, Ph.D. Tel: (303) 496-1550 Email: Robert.Strausberg@nih.gov Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuangui, M.D., Michael Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: David B. Kitzman, Ph.D. DNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center foud through the I.M.A.G.E. Consortium/LNL at: <a href="http://www-bio.lnhi.gov/dbp/Image/Image.html">www-bio.lnhi.gov/dbp/Image/Image.html</a>		
FEATURES	Seq primer: -40m13 fwd. ET from Amersham High quality sequence stop: 367. Location/Qualifiers		
source	1. 489		

[illegible]

RESULT 12  
LOCUS AA225115 491 bp mRNA EST 19-AUG-1997  
DEFINITION nc21g11.s1 NCI\_CGAP\_P11 Homo sapiens CDNA clone IMAGE:100836  
similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);,  
mRNA sequence.

ACCESSION AA225115  
VERSION AA225115.1 GI:1846489  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 491)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

JOURNAL On Apr 14, 1993 this sequence version replaced gi:692766.  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bdrp/image/image.html](http://www.bio.llnl.gov/bdrp/image/image.html)

Insert Length: 1127 Std Error: 0.00  
Seq primer: -41m13 fwd. ET from Amersham  
High quality sequence stop: 335.  
Location/Qualifiers  
1. 491  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="100836"  
/clone\_1ib="NCI\_CGAP\_P11"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/note="Vector: PAMP10; Site\_1: NotI; Site\_2: EcoRI; 1st  
strand CDNA was primed with oligo(dT)17 on 50 ng of  
DNase-treated, total cellular RNA obtained from  
5,000-10,000 microdissected, histologically normal  
prostate epithelial cells. Double-stranded CDNA was  
ligated to EcoRI adaptors, 5 cycles of PCR applied to the  
CDNA with an adaptor-specific primer, and the resulting  
PCR product subcloned into PAMP10 by the UDG-cloning  
method (Life Technologies). Average insert size is 600  
bp. NOTE: Not directionally cloned. This library was  
constructed by David Krizman."

BASE COUNT 121 a 128 c 128 g 111 t 3 others  
ORIGIN

Query Match 98.7%; Score 312; DB 24; Length 491;  
Best Local Similarity 60.8%; Pred. No. 1,08e-49;  
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 55 GAGGAGTCTTGACCCCAAGAACTCAGTGTGACCTCATGTATTCATGAC 114  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|  
QY 1 GARGARTTYTNACNCNAAARAARYTNCARGTGTCAYTNCAYTNATHMSNAYGAY 60  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|

Db 115 GTGTGTGCCCAAGT 128  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|  
QY 61 GTNTGYGCNCARGT 74  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|

RESULT 13  
LOCUS AA574023 505 bp mRNA EST 12-SEP-1997

DEFINITION n144f11.s1 NCI\_CGAP\_P12 Homo sapiens CDNA clone IMAGE:916653  
similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);,  
mRNA sequence.

ACCESSION AA574023  
VERSION AA574023.1 GI:2348538  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 505)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

JOURNAL On Apr 14, 1993 this sequence version replaced gi:692976.  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,  
M.D., Michael Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bdrp/image/image.html](http://www.bio.llnl.gov/bdrp/image/image.html)

Insert Length: 1083 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 348.  
Location/Qualifiers  
1. 505  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="916653"  
/clone\_1ib="NCI\_CGAP\_P12"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/note="Vector: PAMP10; Site\_1: NotI; Site\_2: EcoRI; 1st  
strand CDNA was primed with oligo(dT)17 on 50 ng of  
DNase-treated, total cellular RNA obtained from  
5,000-10,000 microdissected preneoplastic cells  
histologically-determined to be prostatic intraepithelial  
neoplasia 2 (PIN2) cells. Double-stranded CDNA was  
ligated to EcoRI adaptors, 5 cycles of PCR applied to the  
CDNA with an adaptor-specific primer, and the resulting  
PCR product subcloned into PAMP10 by the UDG-cloning  
method (Life Technologies). Average insert size is 600  
bp. NOTE: Not directionally cloned. This library was  
constructed by David Krizman."

BASE COUNT 121 a 124 c 136 g 124 t  
ORIGIN

Query Match 98.7%; Score 312; DB 29; Length 505;  
Best Local Similarity 60.8%; Pred. No. 1,08e-49;  
Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 9 GAGGAGTCTTGACCCCAAGAACTCAGTGTGACCTCATGTATTCATGAC 68  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|  
QY 1 GARGARTTYTNACNCNAAARAARYTNCARGTGTCAYTNCAYTNATHMSNAYGAY 60  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|

Db 69 GTGTGTGCCCAAGT 82  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|  
QY 61 GTNTGYGCNCARGT 74  
|||:||||:| |||:||||:| |||:||||:| |||:||||:| |||:||||:|

RESULT 14  
LOCUS AA534235 520 bp mRNA EST 21-AUG-1997  
DEFINITION n170e05.s1 NCI\_CGAP\_P10 Homo sapiens CDNA clone IMAGE:997856  
similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);,  
mRNA sequence.

ACCESSION AA534235  
 VERSION AA534235.1 GI:2278251  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 520)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 On May 8, 1995 this sequence version replaced gi:801205.  
 CONTACT: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo F. Chuagui,  
 M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: David B. Kitzman, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/URL at:  
[www.bio.lnlnl.gov/dbtrp/image/image.html](http://www.bio.lnlnl.gov/dbtrp/image/image.html)

Insert Length: 1192 Std Error: 0.00  
 Seq primer: -40m13 fwd. Et from Amersham  
 High quality sequence stop: 398.

FEATURES  
 source  
 1..520  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:997856"  
 /clone.lib="NCI\_CGAP\_Pr10"  
 /sex="male"  
 /tissue-type="invasive prostate tumor"  
 /lab\_host="DH10B"  
 /note="Organ: prostate; Vector: pAMP10; mRNA made from  
 invasive prostate tumor, CDNA made by oligo-dt priming.  
 Non-directionally cloned. Size-selected on agarose gel,  
 average insert size 600 bp. Library made by D. Kitzman,  
 NIH."

BASE COUNT 124 a 142 c 139 g 115 t  
 ORIGIN

Query Match 98.7%; Score 312; DB 28; Length 520;  
 Best Local Similarity 60.8%; Pred. No. 1,08e-49;  
 Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 92 GAGGAGTCTTGACCCCAAGAACTTCAGTGTGGACCTCCAGTTATTTCCAAATGAC 151  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 QY 1 GARGARTTYTNMCCNABARARYNCARTGYGTNGAYTNCAYGTATHTWSNAYGAY 60  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 Db 152 GTGTGCGCAAGT 165  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 QY 61 GTNTGYCNCARGT 74

RESULT 15  
 LOCUS AA654548 542 bp mRNA EST 04-NOV-1997  
 DEFINITION nt59111.s1 NCI\_CGAP\_Pr3 Homo sapiens CDNA clone IMAGE:1202805  
 similar to gb:M21895 PROSTATE SPECIFIC ANTIGEN PRECURSOR (HUMAN);,  
 mRNA sequence.  
 ACCESSION AA654548  
 VERSION AA654548.1 GI:2590702  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 542)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL Unpublished (1997)  
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1393439.  
 CONTACT: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,  
 M.D., Michael Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: David B. Kitzman, Ph.D.  
 CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/URL at:  
[www.bio.lnlnl.gov/dbtrp/image/image.html](http://www.bio.lnlnl.gov/dbtrp/image/image.html)

Trace considered overall poor quality  
 Seq primer: -40m13 fwd. Et from Amersham  
 High quality sequence stop: 1.

FEATURES  
 source  
 1..542  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1202805"  
 /clone.lib="NCI\_CGAP\_Pr3"  
 /sex="male"  
 /dev\_stage="45 years old"  
 /lab\_host="DH10B"  
 /note="Vector: pAMP10; Site\_1: NotI; Site\_2: EcoRI; 1st  
 strand cDNA was primed with oligo(dT)17 on 50 ng of  
 DNase-treated, total cellular RNA obtained from  
 5,000-10,000 microdissected cells  
 histologically-determined to be fully malignant prostate  
 cancer cells. Double-stranded cDNA was ligated to EcoRI  
 adaptors, 5 cycles of PCR applied to the cDNA with an  
 adaptor-specific primer, and the resulting PCR product  
 subcloned into pAMP10 by the UDG-cloning method (Life  
 Technologies). Average insert size is 600 bp. NOTE: Not  
 directionally cloned. This library was constructed by  
 David Kitzman."










BASE COUNT 119 a 137 c 167 g 119 t  
 ORIGIN

Query Match 98.7%; Score 312; DB 30; Length 542;  
 Best Local Similarity 60.8%; Pred. No. 1,08e-49;  
 Matches 45; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Db 103 GAGGAGTCTTGACCCCAAGAACTTCAGTGTGGACCTCCAGTTATTTCCAAATGAC 162  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 QY 1 GARGARTTYTNMCCNABARARYNCARTGYGTNGAYTNCAYGTATHTWSNAYGAY 60  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 Db 163 GTGTGCGCAAGT 176  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 QY 61 GTNTGYCNCARGT 74

Search completed: Fri Jun 23 17:10:33 2000  
 Job time : 4615 secs.

Size limit = 25 AA

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

```

msvscn_pp protein - protein database search, using Smith-Waterman algorithm

Run on:      Thu Jun 22 21:27:01 2000;      MasPar time 4.15 Seconds
Tabular output not generated.               142.760 Million cell updates/sec

```

Title:	>US-09-332-866-1
Description:	(1-25) from US09332866.pep
Perfect Score:	183
Sequence:	1 EEFLLPKKLQCVLDHVISNDYCAOV 29

Scoring table: PAM 15  
Gap 15

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Maximum DB seq length 25

```
Database: a-geneseq35
          1: geneseqp
```

statistics: Mean 20.967; Variance 66.122; scale 0.317

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	116	63.4	15	1	W58047	Human prostate specifi	4.36e-005
2	111	60.7	15	1	W58050	Human prostate specifi	1.67e-004
3	110	60.1	15	1	M58048	Human prostate specifi	2.16e-004
4	108	59.0	15	1	M58049	Human prostate specifi	3.73e-004
5	98	53.6	18	1	R65193	Residues 140-157 of th	5.23e-003
6	96	52.5	20	1	W11022	Antigen peptide derive	8.81e-030
7	93	50.8	15	1	M58046	Human prostate specifi	1.92e-020
8	90	49.2	15	1	M58051	Human prostate specifi	4.16e-020
9	90	49.2	17	1	W30784	Prostate Specific Anti	4.16e-020

Note: Post-processor removed 36 summaries from list due to search parameters chosen.

## ALIGNMENTS

RESULT	1
ID	W58047 standard; peptide: 15 AA.
AC	W58047:
DT	11-AUG-1998 (first entry)
DE	Human prostate specific antigen peptide SEQ ID NO:65.
KW	Human; prostate specific antigen; PSA; epitope; monoclonal antibody
KN	detection; cancer; serine protease.

OS Synthetic.  
 OS Homo sapiens.  
 PN M09810282-A1.  
 PF 12-MAR-1998.  
 PF 25-AUG-1997; U194909.  
 PR 06-SEP-1996; US-025404.  
 PR (CENT ) CENTOCOR INC.  
 PI Heaver GA.  
 DR MPI: 98-193789/17.  
 PT Monoclonal antibodies specific for prostate specific antigen -  
 PT useful, e.g. in screening for prostate or breast cancer and  
 PT especially to distinguish between benign prostatic hyperplasia and  
 PT prostate cancer  
 PS Example: Page 56; 84pp; English.  
 CC The present sequence represents a prostate specific antigen (PSA)  
 CC peptide. Monoclonal antibodies specific for PSA and hybridomas producing  
 CC them have been developed. The antibodies: (a) bind to free PSA; (b) are  
 CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKRFLL  
 CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and  
 CC 15.2, 156 or 225 binding to amino acids 139-144 (ELFLFLLP) and 55-60  
 CC (SLRHP). respectively of free and bound PSA, or fragments. The antibodies  
 CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
 CC comprising a solid support with attached labelled monoclonal antibody  
 CC specific for free PSA (especially (b)) and a PSA standard can be used.  
 CC To detect both free and bound PSA, a second solid support with attached  
 CC (differently labelled) monoclonal antibody binding free and bound PSA  
 CC (especially selected from (c)) can be used either with, or in place of,  
 CC the first solid support. The antibodies are useful in cancer screening,  
 CC especially prostate and breast cancer. By obtaining total and free PSA  
 CC values, their ratio can be used to separate prostatic cancer (PCA) from  
 CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
 CC useful after radical prostatectomy, to predict disease persistence. The  
 CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
 CC possible by total PSA testing, avoiding biopsies.  
 SQ Sequence 15 AA;

Query Match	63.4%	Score 116;	DB 1;	Length 15.
Best Local Similarity	100.0%;	Pred. No. 4.36e-05;		
Matches	15;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;

D6	1	EEFLTPKKLQCVDLH	15
QY	1	EEFLTPKKLQCVDLH	15

## RESULT 2

ID	W58048	standard; peptide; 15 AA.
AC	W58048:	

DT	11-AUG-1998 (first entry)
DE	Human contact

Human; prostate specific antigen: PSA: antipeptide: prostate specific antigen peptide seq ID NO: 66

synthetic detection; cancer; serine protease.

*Homo sapiens.*

PD 12-MAR-1998

PF ~~25-AUG-1997~~; U14909  
DB 06-07-1997

PA (CENZ ) CENTOCOR INC

P1 Heavenr GA;  
DB WPT: 08-103700/17

PT Monoclonal antibodies

especially to distribute

PT	prostate cancer	prostate cancer
PC	prostate cancer	prostate cancer

The present sequence represents

peptide. Monoclonal antibodies specific for PSA and hybridomas producing them have been developed. The

monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LYNNE

15.2, 156 or 225 binding to amino acids 138-144 (entry no. CC

CC (SLRPE) respectively of free and bound PSA, or fragments. The antibodies  
CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
CC comprising a solid support with attached labelled monoclonal antibody  
CC specific for free PSA (especially (b)) and a PSA standard can be used.  
CC To detect both free and bound PSA, a second solid support with attached  
CC (differently labelled) monoclonal antibody binding free and bound PSA  
CC (especially selected from (c)) can be used either with, or in place of,  
CC the first solid support. The antibodies are useful in cancer screening,  
CC especially prostate and breast cancer. By obtaining total and free PSA  
CC values, their ratio can be used to separate prostatic cancer (PCA) from  
CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
CC useful after radical prostatectomy, to predict disease persistence. The  
CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
CC possible by total PSA testing, avoiding biopsies.

SO Sequence 15 AA:  
Query Match 60.7%; Score 111; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.67e-04;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 LTPKLCQVDLHVIS 15  
Qy 4 LTPKLCQVDLHVIS 18

## RESULT

3  
ID W58050 standard; peptide: 15 AA.  
AC W58050.  
DT 11-AUG-1998 (first entry)  
DE Human prostate specific antigen peptide SEQ ID NO:68.  
KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
KW detection; cancer; serine protease.  
OS Synthetic.  
OS Homo sapiens.  
PN W09810292-AL.  
PD 12-MAR-1998.  
PF 25-AUG-1997; U14909.  
PR 06-SEP-1996; US-025404.  
PA (GENZ) CENTOCOR INC.  
PI Heavner GA.  
DR WPI: 98-193789/17.  
PT Monoclonal antibodies specific for prostate specific antigen -  
PT useful, e.g. in screening for prostate or breast cancer and  
PT especially to distinguish between benign prostatic hyperplasia and  
PT prostate cancer.  
PS Example: Page 58; 84pp; English.  
CC The present sequence represents a prostate specific antigen (PSA)  
CC peptide. Monoclonal antibodies specific for PSA and hybridomas producing  
CC them have been developed. The antibodies: (a) bind to free PSA; (b) are  
CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL)  
CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22-2 and  
CC 15-2, 156 or 225 binding to amino acids 139-144 (EELFRT) and 55-60  
CC (SLRPE) respectively of free and bound PSA, or fragments. The antibodies  
CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
CC comprising a solid support with attached labelled monoclonal antibody  
CC specific for free PSA (especially (b)) and a PSA standard can be used.  
CC To detect both free and bound PSA, a second solid support with attached  
CC (differently labelled) monoclonal antibody binding free and bound PSA  
CC (especially selected from (c)) can be used either with, or in place of,  
CC the first solid support. The antibodies are useful in cancer screening,  
CC especially prostate and breast cancer. By obtaining total and free PSA  
CC values, their ratio can be used to separate prostatic cancer (PCA) from  
CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
CC useful after radical prostatectomy, to predict disease persistence. The  
CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
CC possible by total PSA testing, avoiding biopsies.

SO Sequence 15 AA:  
Query Match 60.1%; Score 110; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.19e-04;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 QCVDLHVISNDVCAQ 15  
Qy 10 QCVDLHVISNDVCAQ 24

## RESULT

4  
ID W58049 standard; peptide: 15 AA.  
AC W58049.  
DT 11-AUG-1998 (first entry)  
DE Human prostate specific antigen peptide SEQ ID NO:67.  
KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
KW detection; cancer; serine protease.  
OS Synthetic.  
OS Homo sapiens.  
PN W09810292-AL.  
PD 12-MAR-1998.  
PF 25-AUG-1997; U14909.  
PR 06-SEP-1996; US-025404.  
PA (GENZ) CENTOCOR INC.  
PI Heavner GA.  
DR WPI: 98-193789/17.  
PT Monoclonal antibodies specific for prostate specific antigen -  
PT useful, e.g. in screening for prostate or breast cancer and  
PT especially to distinguish between benign prostatic hyperplasia and  
PT prostate cancer.  
PS Example: Page 57; 84pp; English.  
CC The present sequence represents a prostate specific antigen (PSA)  
CC peptide. Monoclonal antibodies specific for PSA and hybridomas producing  
CC them have been developed. The antibodies: (a) bind to free PSA; (b) are  
CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL)  
CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22-2 and  
CC 15-2, 156 or 225 binding to amino acids 139-144 (EELFRT) and 55-60  
CC (SLRPE) respectively of free and bound PSA, or fragments. The antibodies  
CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
CC comprising a solid support with attached labelled monoclonal antibody  
CC specific for free PSA (especially (b)) and a PSA standard can be used.  
CC To detect both free and bound PSA, a second solid support with attached  
CC (differently labelled) monoclonal antibody binding free and bound PSA  
CC (especially selected from (c)) can be used either with, or in place of,  
CC the first solid support. The antibodies are useful in cancer screening,  
CC especially prostate and breast cancer. By obtaining total and free PSA  
CC values, their ratio can be used to separate prostatic cancer (PCA) from  
CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
CC useful after radical prostatectomy, to predict disease persistence. The  
CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
CC possible by total PSA testing, avoiding biopsies.

SO Sequence 15 AA:  
Query Match 59.0%; Score 108; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3.73e-04;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KKLQCVDLHVISNDV 15  
Qy 7 KKLQCVDLHVISNDV 21

## RESULT

5  
ID W65193 standard; peptide: 18 AA.  
AC W65193.  
DT 24-AUG-1995 (first entry)  
DE Residues 140-157 of the deduced sequence of human kallikrein hK2.  
KW Prostate specific glandular kallikrein; hK2; prostate cancer.  
OS Homo sapiens.  
PN W09503334-A.  
PD 02-FEB-1995.  
PF 28-JUN-1994; U07329.  
PR 22-JUL-1993; US-096946.  
PA (MAYO-) MAYO FOUNDATION.  
PI Klee GG, Tindall DJ, Young CYF, Young CY.  
DR WPI: 95-075194/10.

PT New antibodies specific for human prostate glandular kallikrein  
 PT used for purifications and in assays which can be used for the  
 PT detection and staging of prostatic cancer  
 CC disclosure: Page 23; 43pp: English.  
 CC A polypeptide of the formula in R65190 is claimed. It represents  
 CC residues 41-56 of the deduced AA sequence of human prostate specific  
 CC glandular kallikrein (hk2) as reported by L.J. Schedlish et al., DNA,  
 CC 6, 429, 1987. Immunogenic or immunoreactive subunits of this peptide,  
 CC pref. of at least peptidyl units, which can be specifically bound  
 CC by specific anti-hk2 antibodies are also within the scope of the  
 CC invention. Other immunogenic hk2 subunits will also generate an  
 CC antibody response that is unique to hk2, including the sequences in  
 CC R65191-R65198. The numbering of the AAs identify the hk2 residues  
 CC disclosed in Schedlish et al., cited above.  
 SQ Sequence 18 AA;

DB Query Match 53.68; Score 96; DB 1; Length 18;  
 Best Local Similarity 66.78; Pred. NO. 5.23e-03;  
 Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

DB 1 EFLPRSLQCVSLHLN 18  
 2 EFLPKKLCQVDLHVSN 19

RESULT 6  
 ID W11022 standard; peptide: 20 AA.

AC W11022;  
 DT 27-OCT-1997 (first entry)  
 DE Antigen peptide derived from human prostate specific antigen.  
 KM PSA; kallikrein: prostate cancer; antigenic; polyclonal antisera.  
 OS Homo sapiens.  
 PN W09640754-A1.

PF 06-JUN-1996; U09303.  
 PR 07-JUN-1995; US-472228.  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PI Fritsche HA, Johnston DA, Kokolus WJ;  
 DR WPI; 97-108633/10.  
 PT Antigen peptide derived from prostate specific antigen - does not  
 PT cross react with related kallikreins, for diagnosis of prostate  
 PT cancer

PS Example 2; Page 42; 74pp: English.

CC The present sequence represents a novel peptide which has 20 contiguous  
 CC amino acids derived from the 240 residue sequence of the human prostate  
 CC specific antigen (PSA) (see W11023). The preferred peptide has two  
 CC hydrophobic regions and one hydrophilic region each of about 5 aa in  
 CC length, arranged as follows: hydrophobic-hydrophilic-hydrophobic.  
 CC Polyclonal antisera specific for the antigenic peptide may be used in a  
 CC method for diagnosing prostate cancer in vivo or in vitro. The peptide  
 CC represents a sequence unique to PSA which does not cross react with  
 CC certain kallikreins.  
 SQ Sequence 20 AA;

Query Match 52.58; Score 96; DB 1; Length 20;  
 Best Local Similarity 92.94; Pred. NO. 8.81e-03;  
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 7 KRLQCVLHVSN 20  
 7 KRLQCVLHVSN 20

RESULT 7  
 ID W58046 standard; peptide: 15 AA.

AC W58046;  
 DT 11-AUG-1998 (first entry)  
 DE Human prostate specific antigen peptide SEQ ID NO:64.  
 KM Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
 KM detection; cancer; serine protease.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09810292-A1.

PD 12-MAR-1998.  
 PR 25-AUG-1997; U14909.  
 PR 06-SEP-1996; US-025404.  
 PA (CENZ ) CENTOCOR INC.  
 PI Heavner GA;  
 DR WPI; 98-193789/17.

PT Monoclonal antibodies specific for prostate specific antigen -  
 PT useful, e.g. in screening for prostate or breast cancer and  
 PT especially to distinguish between benign prostatic hyperplasia and  
 PT prostate cancer  
 PS Example; Page 36; 84pp: English.

CC The present sequence represents a prostate specific antigen (PSA)  
 CC peptide. Monoclonal antibodies specific for PSA and hydridomas producing  
 CC them have been developed. The antibodies: (a) bind to free PSA; (b) are  
 CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFU).  
 CC 15.2, 156 or 225 binding to amino acids 139-144 (EFLFLNP) and 55-60  
 CC (SLFHPD) respectively of free and bound PSA, or fragments. The antibodies  
 CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
 CC comprising a solid support with attached labelled monoclonal antibody  
 CC specific for free PSA (especially (b)) and a PSA standard can be used.  
 CC To detect both free and bound PSA, a second solid support with attached  
 CC (differently labelled) monoclonal antibody binding free and bound PSA  
 CC (especially selected from (c)) can be used either with, or in place of,  
 CC the first solid support. The antibodies are useful in cancer screening,  
 CC especially prostate and breast cancer. By obtaining total and free PSA  
 CC values, their ratio can be used to separate prostatic cancer (PCA) from  
 CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also  
 CC useful after radical prostatectomy, to predict disease persistence. The  
 CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
 CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
 CC possible by total PSA testing, avoiding biopsies.  
 SQ Sequence 15 AA;

Query Match 50.88; Score 93; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. NO. 1.92e-02;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 4 EFLPKKLCQV 15  
 1 EFLPKKLCQV 12

RESULT 8  
 ID W58051 standard; peptide: 15 AA.

AC W58051;  
 DT 11-AUG-1998 (first entry)  
 DE Human prostate specific antigen peptide SEQ ID NO:69.  
 KM Human; prostate specific antigen; PSA; epitope; monoclonal antibody;  
 KM detection; cancer; serine protease.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09810292-A1.

PD 12-MAR-1998.  
 PR 25-AUG-1997; U14909.  
 PR 06-SEP-1996; US-025404.  
 PA (CENZ ) CENTOCOR INC.  
 PI Heavner GA;  
 DR WPI; 98-193789/17.

PT Monoclonal antibodies specific for prostate specific antigen -  
 PT useful, e.g. in screening for prostate or breast cancer and  
 PT especially to distinguish between benign prostatic hyperplasia and  
 PT prostate cancer  
 PS Example; Page 58; 84pp: English.

CC The present sequence represents a prostate specific antigen (PSA)  
 CC peptide. Monoclonal antibodies specific for PSA and hydridomas producing  
 CC them have been developed. The antibodies: (a) bind to free PSA; (b) are  
 CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFU).  
 CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and  
 CC 15.2, 156 or 225 binding to amino acids 139-144 (EFLFLNP) and 55-60  
 CC (SLFHPD) respectively of free and bound PSA, or fragments. The antibodies  
 CC are useful to detect PSA. For detecting free PSA only, an immunoassay  
 CC comprising a solid support with attached labelled monoclonal antibody

CC specific for free PSA (especially (b)) and a PSA standard can be used.  
 CC To detect both free and bound PSA, a second solid support with attached  
 CC (differently labeled) monoclonal antibody binding free and bound PSA  
 CC (especially selected from (c)) can be used either with, or in place of,  
 CC the first solid support. The antibodies are useful in cancer screening,  
 CC especially prostate and breast cancer. By obtaining total and free PSA  
 CC values, their ratio can be used to separate prostatic cancer (PCA) from  
 CC benign prostatic hyperplasia (BPH) patients. Measurement of PSA is also  
 CC useful after radical prostatectomy, to predict disease persistence. The  
 CC antibodies allow evaluation of PSA free/total ratio, enabling separation  
 CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously  
 CC possible by total PSA testing, avoiding biopsies.  
 SQ Sequence 15 AA;

Query Match 49.2%; Score 90; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 4.16e-02;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 DLHVISNDVCAQY 13  
 |||||  
 QY 13 DLHVISNDVCAQY 25

RESULT 9  
 ID W30784 standard; peptide: 17 AA.

AC W30784;  
 DT 27-FEB-1998 (first entry)  
 DE Prostate Specific Antigen (PSA) derived peptide ABR16.  
 KW Prostate specific antigen; PSA; immunogenic; human glandular kallikrein;  
 KW HK2; antigen; antibody; detection; diagnosis; prostate cancer.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN WO9729199-A2.  
 PD 14-AUG-1997.  
 PF 06-FEB-1997; U01911.  
 PR 06-FEB-1996; US-595945.  
 PA (ABBO) ABBOTT LAB.  
 PI Bridon DP, Dowell BL, Lilja H, Pettersson IK, Pihonen T;  
 PI Qiu X, Vihtinen MA;  
 DR WPI: 97-415352/38.  
 PT Prostate Specific Antigen peptide(s) - useful for diagnosis of  
 PT prostate cancer  
 PS Claim 2, Page 10, 42pp; English.  
 CC W30769-84 are synthetic peptides derived from the prostate specific  
 CC antigen (PSA) sequence. These peptides are identical to a highly  
 CC immunogenic region of PSA, and also comprise one or more amino acids  
 CC identical or non-identical to the amino acid sequence of human glandular  
 CC kallikrein (HK2). The peptides are used as antigens for the production of  
 CC antibodies which are used to detect PSA in a test sample (claimed). This  
 CC is useful for diagnosis of prostate cancer. The peptides enable the  
 CC production of antisera necessary to determine the amount of total PSA,  
 CC free PSA and PSA-ACP complex present in a sample and thus improve the  
 CC ability of the clinician to distinguish, e.g., between BPH (benign  
 CC prostatic hyperplasia) and prostatic cancer in a patient.  
 SQ Sequence 17 AA;

Query Match 49.2%; Score 90; DB 1; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 4.16e-02;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 DLHVISNDVCAQY 13  
 |||||  
 QY 13 DLHVISNDVCAQY 25

Search completed: Thu Jun 22 21:27:06 2000  
 Job time : 5 secs.



\*\*\*\*\*  
 RELEASE  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Mpsrch\_dp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Thu Jun 22 21:24:52 2000; Maspar time 3.43 Seconds  
 Tabular output not generated. 105.343 Million cell updates/sec

Title: >US-09-332-866-1  
 Description: (1-25) from US09332866.pep  
 Perfect Score: 183  
 Sequence: 1 EFLTPKRLQCVDLHVISNDVCAQV 25

Scoring table:  
 Gap 15  
 PAM 150

Searched: 145341 seqs, 14437480 residues

Post-processing: Minimum Match 08  
 Listing first 45 summaries  
 Maximum DB seq length 25

Database: a-issued  
 1:5h\_COMB 2:5b\_COMB 3:6\_COMB 4:PCT\_COMB 5:backfiles1

Statistics: Mean 19.756; Variance 64.432; scale 0.307

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	98	53.6	18	4	PCT-US94-0 Sequence 4, Applicatio	3.47e-03
2	98	53.6	18	1	US-08-096- Sequence 4, Applicatio	3.47e-03
3	96	52.5	20	1	US-08-472- Sequence 22, Applicati	5.79e-03
4	96	52.5	20	4	PCT-US96-0 Sequence 22, Applicati	5.79e-03

Note: Post-processor removed 41 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1  
 ID PCT-US94-07329-4 STANDARD: PRT: 18 AA.  
 AC xxxxxx  
 XX  
 XX  
 DT  
 DE Sequence 4, Application PC/TUS9407329  
 CC Sequence 4, Application PC/TUS9407329  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Mayo Foundation for Medical

CC APPLICANT: Education and Research  
 CC TITLE OF INVENTION: Antibodies Specific for Human  
 CC TITLE OF INVENTION: Prostate Glandular kallikrein  
 CC NUMBER OF SEQUENCES: 11  
 CC CORRESPONDENCE ADDRESSES:  
 CC ADDRESSEE: Schweigman, Lundberg & Woessner  
 CC STREET: 3500 IDS Center  
 CC CITY: Minneapolis  
 CC STATE: MN  
 CC COUNTRY: USA  
 CC ZIP: 55402  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: PCT/US94/07329  
 CC FILING DATE:  
 CC CLASSIFICATION:  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Raasch, Kevin W.  
 CC REGISTRATION NUMBER: 35,651  
 CC REFERENCE/DOCKET NUMBER: 150.62MO  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 612-339-0331  
 CC TELEFAX: 612-339-3061  
 CC INFORMATION FOR SEQ ID NO: 4:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 18 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: peptide  
 CC IMMEDIATE SOURCE:  
 CC CLONE: hK2 amino acid residues 140-157  
 CC SEQUENCE 18 AA: 2112 MW: 1815 CN;  
 CC  
 Db 1 EFLRPSLQCVSLHLISN 18  
 ||| | : ||| | : ||| |  
 Qy 2 EFLTPKRLQCVDLHVISN 19  
 Query Match 53.6%; Score 98; DB 4; Length 18;  
 Best Local Similarity 66.7%; Pred. No. 3.47e-03;  
 Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 DE Sequence 4, Application US/08096946  
 XX  
 XX  
 AC xxxxxx  
 XX  
 XX  
 DT  
 DE Sequence 4, Application US/08096946  
 CC Patent No. 5516639  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Tindall, Donald J.  
 CC APPLICANT: Young, Charles Y-F  
 CC APPLICANT: Klee, George G.  
 CC TITLE OF INVENTION: Antibodies Specific for Human  
 CC TITLE OF INVENTION: Prostate Glandular kallikrein  
 CC NUMBER OF SEQUENCES: 11  
 CC CORRESPONDENCE ADDRESSES:  
 CC ADDRESSEE: Merchant & Gould  
 CC STREET: 3100 No. 5516639west Center  
 CC CITY: Minneapolis  
 CC STATE: MN  
 CC COUNTRY: USA  
 CC ZIP: 55402



\*\*\*\*\*  
 MIPERLH  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Msrch<sub>LP</sub> protein - protein database search, using Smith-Waterman algorithm  
 Run on: Thu Jun 22 21:26:38 2000; Maspar time 5.78 Seconds  
 Tabular output not generated. 203.949 Million cell updates/sec

Title: >US-09-332-866-1  
 Description: (1-25) from US09332866.pep  
 Perfect Score: 183  
 Sequence: 1 EEFLTPKKLQCVDLHVISNDVCAQV 25

Scoring table:  
 PAM 150  
 Gap 15

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 08  
 Listing first 45 summaries  
 Maximum DB seq length 25

Database: p1r62

1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 28.844; Variance 40.415; scale 0.714

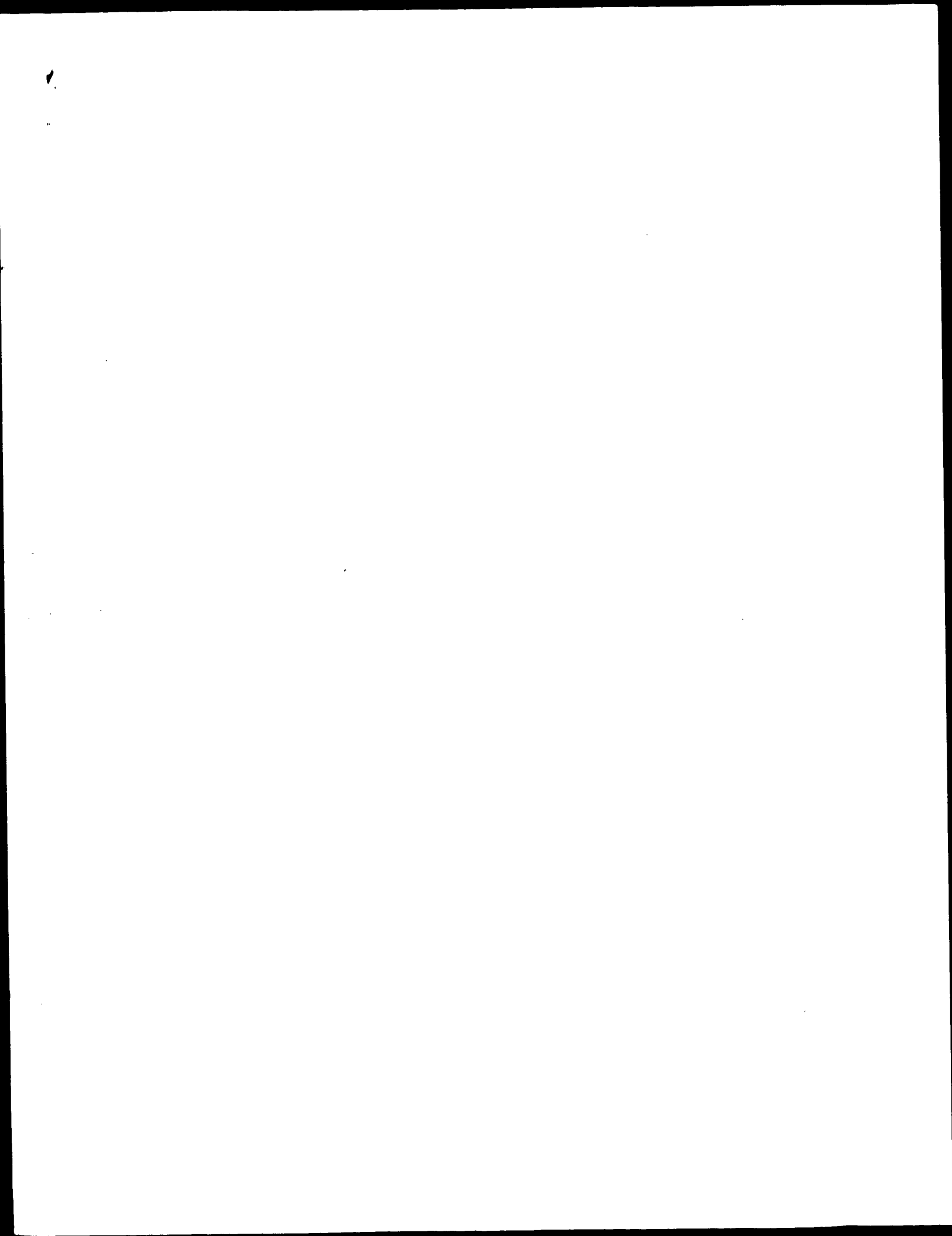
Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.
*****					

No matches found.

Search completed: Thu Jun 22 21:26:45 2000  
 Job time : 7 secs.



\*\*\*\*\*  
 MIPSEK  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Mpsrch.p protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 22 21:25:48 2000; MasPar time 3.90 Seconds

Tabular output not generated. 195.186 Million cell updates/sec

Title: >US-09-332-866-1  
 Description: (1-25) from US09332866.pep  
 Perfect Score: 183  
 Sequence: 1 EEFLPKKLCQVDLHVISNDVCAQV 25

Scoring table: PAM 150  
 Gap 15

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries  
 Maximum DB seq length 25

Database: swiss-prot38  
 1:swissprot

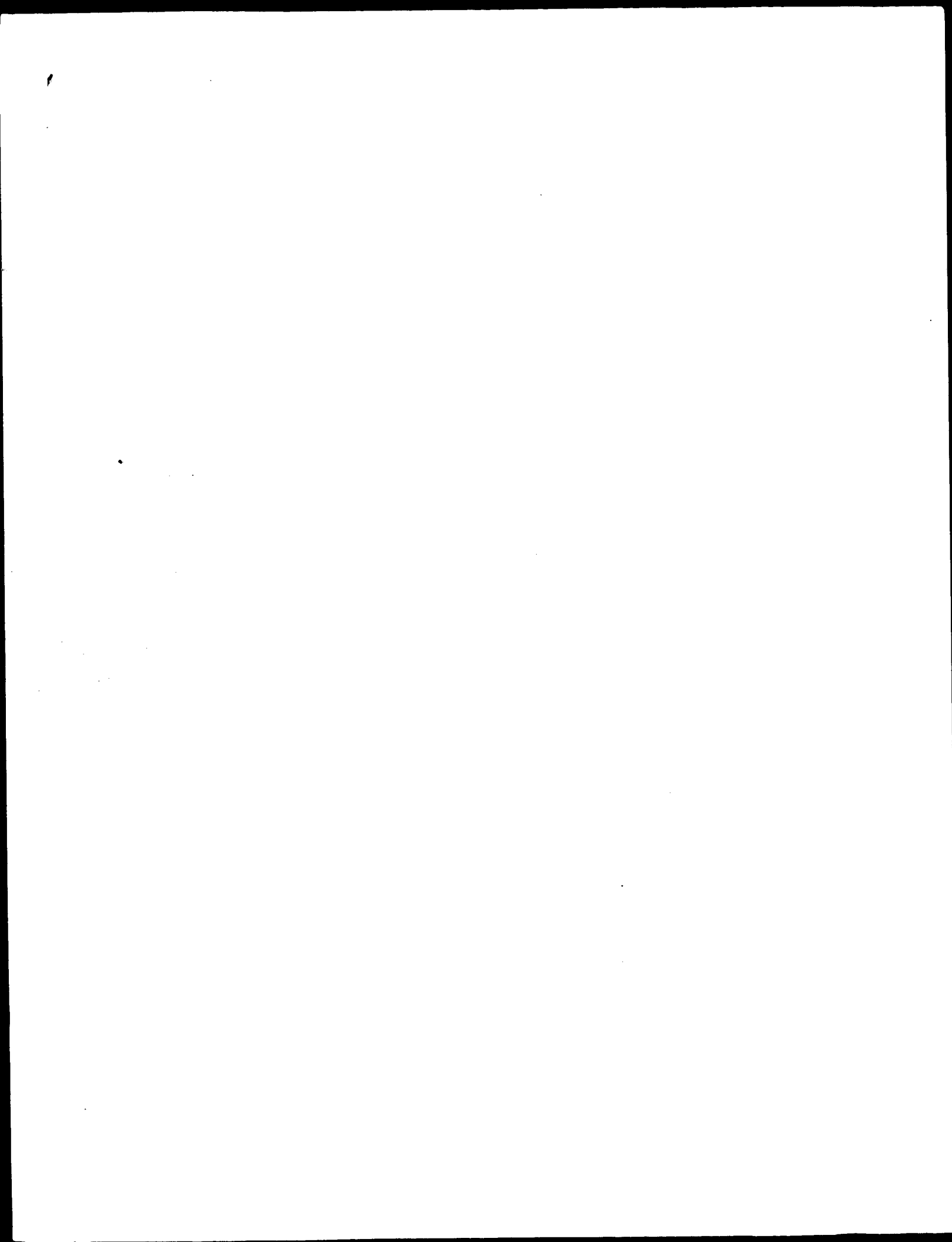
Statistics: Mean 29.602; Variance 37.132; scale 0.797

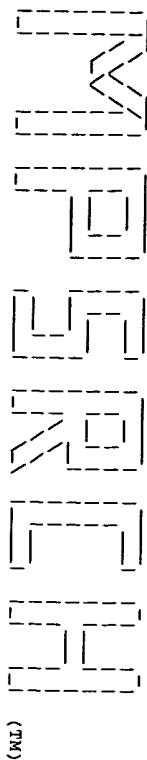
Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.
No matches found.					

Search completed: Thu Jun 22 21:25:53 2000  
 Job time : 5 secs.





Release 3.1a John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mparch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 22 21:26:11 2000; Maspar time 9.41 Seconds 184.120 Million cell updates/sec

Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 183  
Sequence: 1 EEFLTPKKLQCVDLHVISNDVCAQV 25

Scoring table:  
PAM 150  
Gap 15

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Maximum DB seq length 25

Database: sptrembl12  
1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
9:sp\_phase 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 28.558; Variance 35.870; scale 0.796

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
------------	-------	-------------	--------	----	-------------	-----------

No matches found.

Search completed: Thu Jun 22 21:26:22 2000  
Job time : 11 secs.





Transl. Seq 1 AA → 104 Size = 25

W T E S R L H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:30:31 2000; MasPar time 285.10 Seconds  
1121.254 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTNACNCCNAA.....AYGAYGNTGYGCNCARGTN 75  
Comp: CTCTCTAARARANTGNGGNTT.....TRCTRCANACRCGNGTTCAN

Scoring table: TABLE bktranslated2  
gap 40

Mismatch STD : Dbase 0; Query 0

Searched: 23132 seqs, 2311144889 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Maximum DB seq length 25

Database: emb161  
1:em\_ba1 2:em\_ba2 3:em\_fun 4:em\_htg1 5:em\_htg2 6:em\_htg3  
7:em\_hum1 8:em\_hum2 9:em\_hum3 10:em\_hum4 11:em\_in  
12:em\_om 13:em\_or 14:em\_ov 15:em\_pat 16:em\_pl 17:em\_ro  
18:em\_sts 19:em\_un 20:em\_v1  
genbank116  
21:gb\_htg1 22:gb\_htg10 23:gb\_htg11 24:gb\_htg12  
25:gb\_htg13 26:gb\_htg14 27:gb\_htg2 28:gb\_htg3 29:gb\_htg4  
30:gb\_htg5 31:gb\_htg6 32:gb\_htg7 33:gb\_htg8 34:gb\_htg9

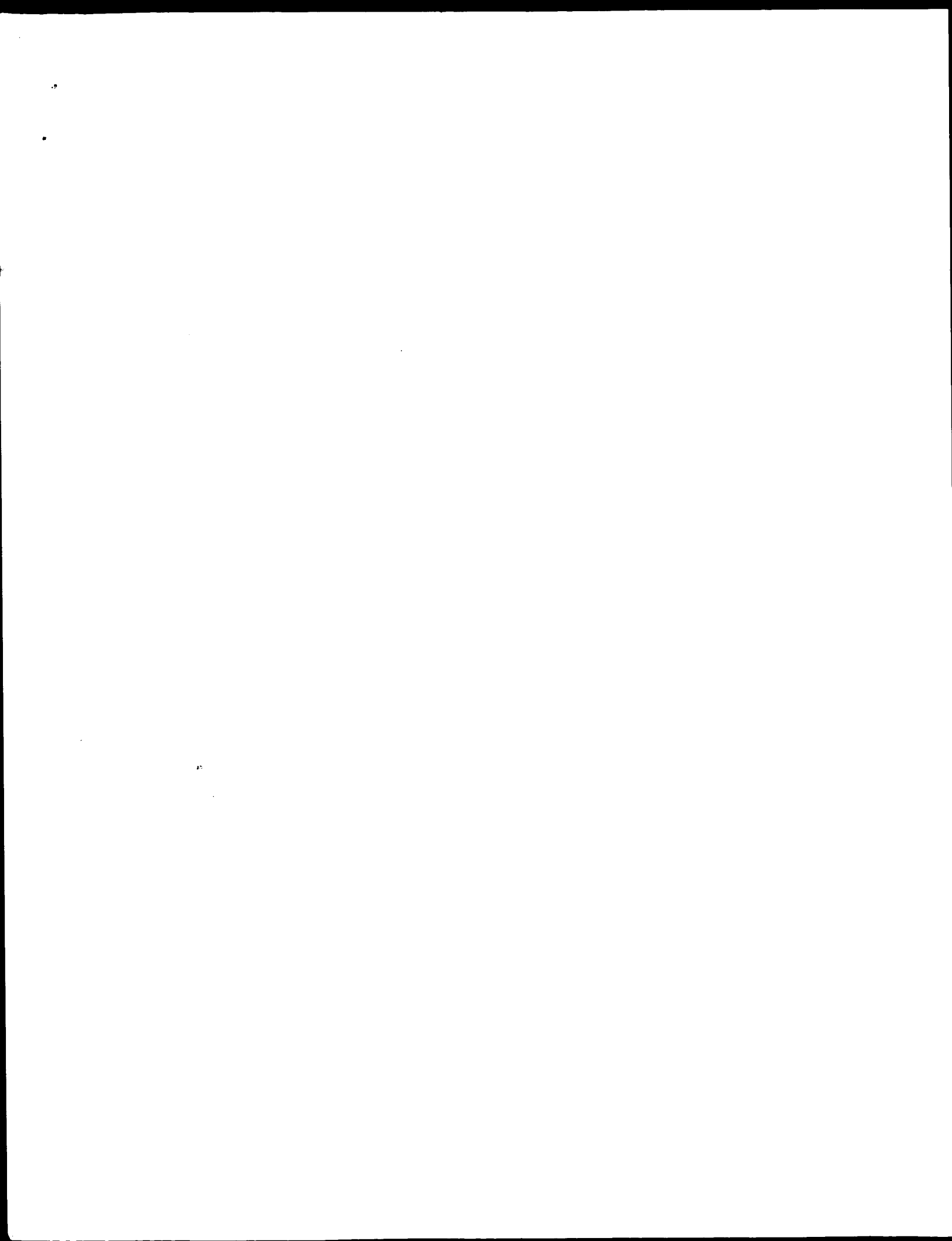
Statistics: Mean 78.372; Variance 327.890; scale 0.239

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
No matches found.						

Search completed: Fri Jun 23 17:35:29 2000  
Job time : 298 secs.



\*\*\*\*\*

NIPOSEIL (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Msrch\_tpn n.a. n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:24:02 2000; MasPar time 177.49 Seconds  
Tabular output not generated. 1034.812 Million cell updates/sec

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GANGARTTYTNACNCNAA.....AYGATGTGCGNCARGTN 75  
Comp: CTYCTVARRANGNGNT.....TRCTRCANACRCNGGTTCAN

Scoring table: TABLE dktranslate2  
Gap 40

Nmatch STD: Dbase 0; Query 0

Searched: 667327 seqs, 1224492533 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Maximum DB seq length 25

Database: genbank2-116  
1:gb\_pat 2:gb\_ph 3:gb\_pl1 4:gb\_pl2 5:gb\_pl3 6:gb\_pr1  
7:gb\_pr2 8:gb\_pr3 9:gb\_pr4 10:gb\_pr5 11:gb\_ro 12:gb\_sts  
13:gb\_sy 14:gb\_un 15:gb\_v1

Statistics: Mean 49.932; Variance 184.198; scale 0.271

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description	Pred. No.
No matches found.						

Search completed: Fri Jun 23 17:27:04 2000  
Job time : 182 secs.



RELEASE

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_tpm n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 18:53:33 2000; Maspar time 22.92 Seconds  
791.884 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTNACNCNNA.....AYGAYGTNTGYGCNCARGTN 75  
Comp: CTYCTYAARRANTGNGNGNT.....TRCTRCANANCRGNGNTYCAN

Scoring table: TABLE bkttranslate2  
Gap 40

Mmatch STD : Dbase 0; Query 0

Searched: 311545 seqs, 121019393 bases x 2

Post-processing: Minimum Match 08  
Listing first 45 summaries  
Maximum DB seq length 25

Database: n-geneseq36  
1:geneseqn

Statistics: Mean 40.189; Variance 176.566; scale 0.228

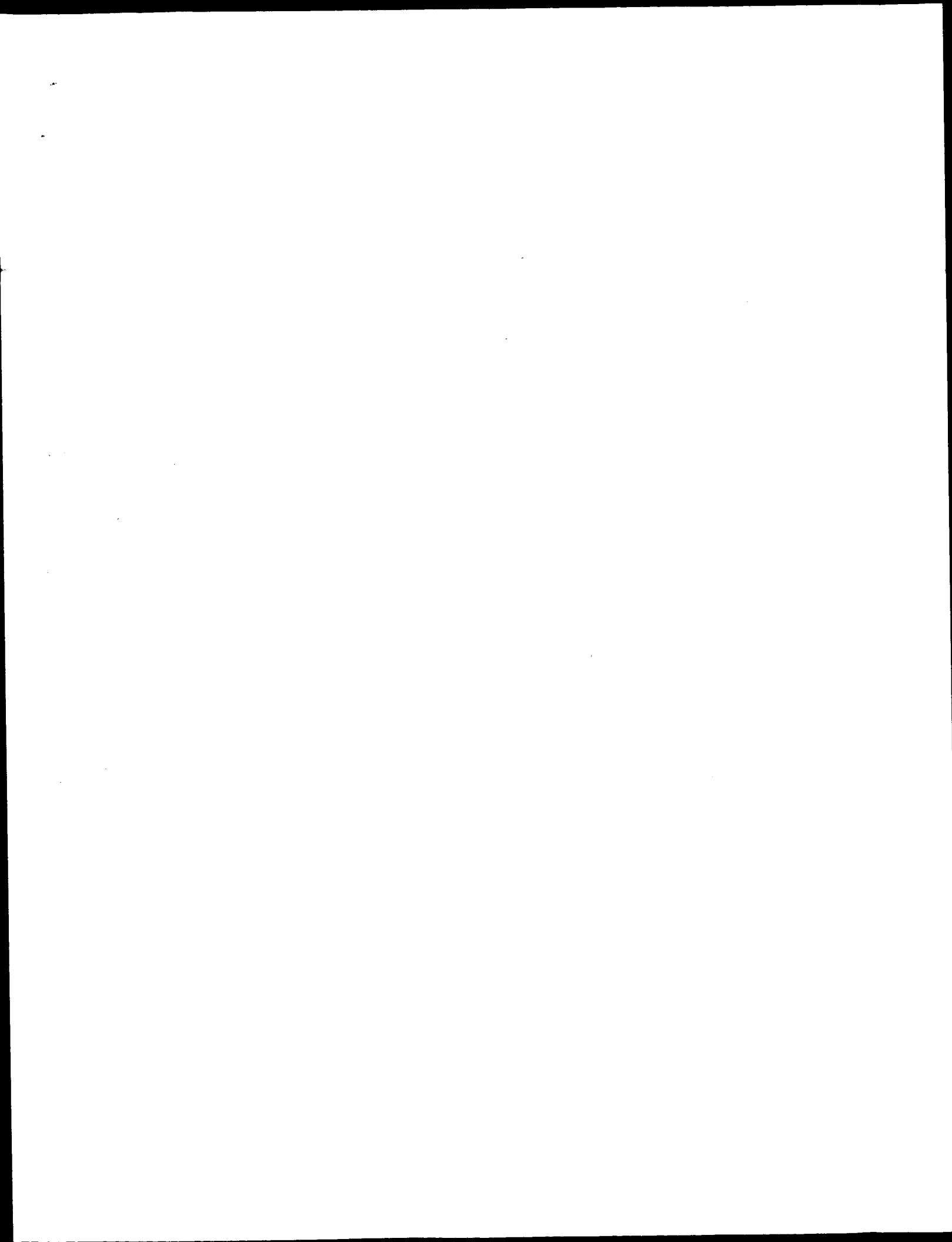
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result	Query	No.	Score	Match	Length	DB	ID	Description	Pred.	No.
--------	-------	-----	-------	-------	--------	----	----	-------------	-------	-----

No matches found.

Search completed: Fri Jun 23 18:54:00 2000  
Job time : 27 secs.



WIPSEH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 18:54:17 2000; Maspar time 15.21 Seconds  
Tabular output not generated. 640.747 Million cell updates/sec

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTNACNCNA.....AYGATGNTGYGNCARGTN 75  
Comp: CTYCTVAAARANTGNGSNT.....TRCTRCANACRCGNGTYCAN

Scoring table: TABLE dkttranslate2  
Gap 40

Nmatch STD : Dbase 0; Query 0

Searched: 230473 seqs, 64993002 bases x 2

Post-processing: Minimum Match 08  
Listing first 45 summaries  
Maximum DB seq length 25

Database: n-Issued  
1:5A\_COMB 2:5B\_COMB 3:5C\_COMB 4:5D\_COMB 5:6\_COMB  
6:PCT\_COMB 7:backfiles1

Statistics: Mean 37.556; Variance 145.681; scale 0.258

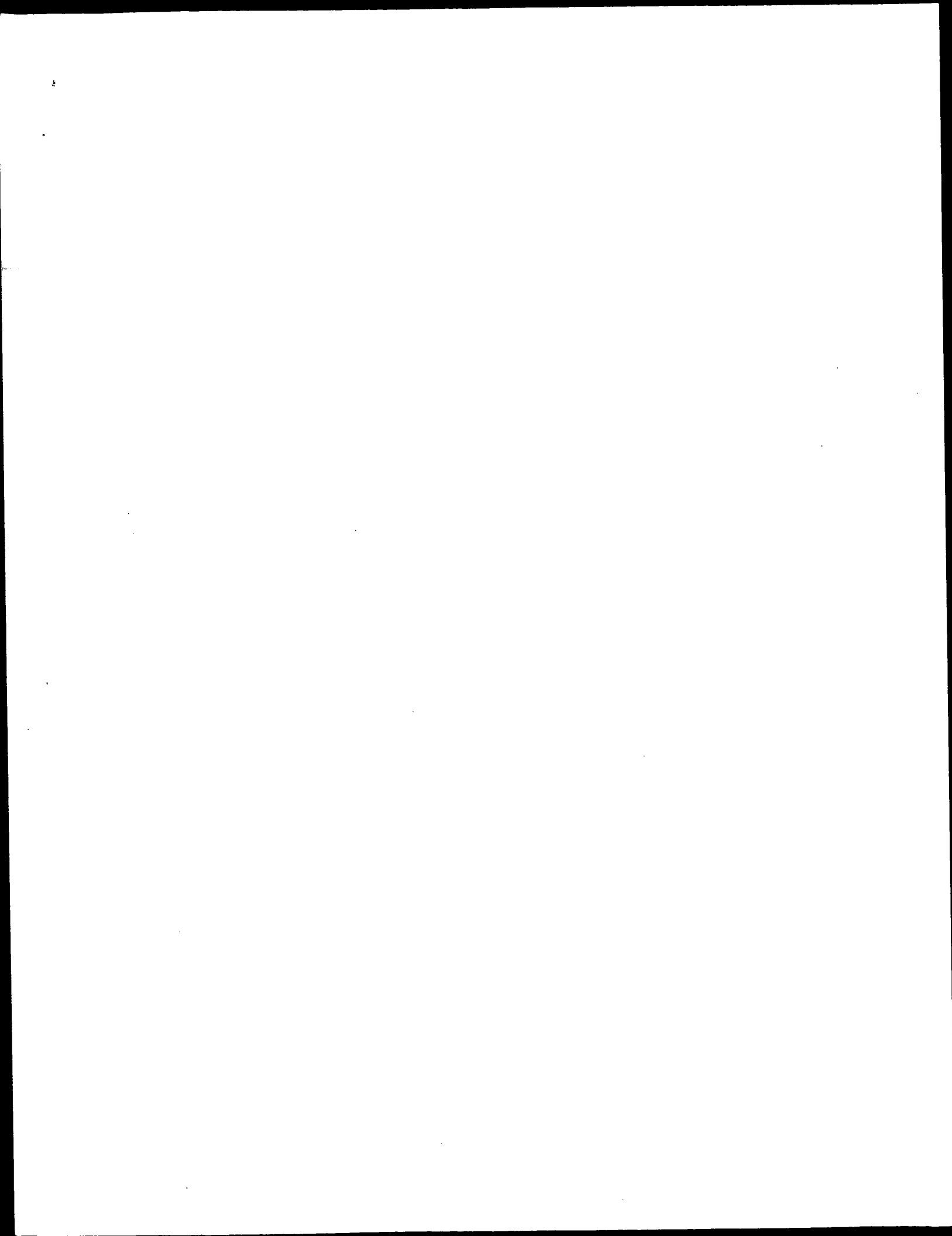
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match length	ID	Description	Pred. No.
.....				

No matches found.

Search completed: Fri Jun 23 18:54:36 2000  
Job time : 19 secs.







Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 18:54:56 2000; MasPar time 248.75 Seconds  
809.968 Million cell updates/sec

Tabular output not generated.

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GARGARTTYTNACNCNAA.....AYGAYGNTNGCGNCARGTN 75  
Comp: CTCTTAAARANTNGNGNTT.....TRCTRCANACRCGNGTYCAN

Scoring table: TABLE bkttranslated  
Gap 40

Nmatch SMD : Dbase 0; Query 0

Searched: 3601417 seqs, 1343217448 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Maximum DB seq length 25

Database: n-pending1  
1:U6 2:U7 3:U80 4:U81A 5:U81B 6:U81C 7:U81D 8:U82A  
9:U82B 10:U82C 11:U83A 12:U83B 13:U84A 14:U84B 15:U84C  
16:U85 17:U86 18:U87A 19:U87B 20:U87C 21:U88A 22:U88B  
23:U88C 24:U88D 25:U89A 26:U89B 27:U89C 28:U89D 29:U89E  
30:U89F 31:U89G 32:U89H 33:U89I 34:U90A 35:U90B 36:U90C  
37:U90D 38:U90E 39:U91A 40:U91B 41:U91C 42:U92A 43:U92B  
44:U92C 45:U92D 46:U92E 47:U93A 48:U93B 49:U93C 50:U93D  
51:U93E 52:U93F 53:U94A 54:U94B 55:U94C 56:U95 57:NEW08  
58:NEW09

Statistics: Mean 49.471; Variance 110.484; scale 0.448

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.
No matches found.					

Search completed: Fri Jun 23 18:59:11 2000  
Job time : 255 secs.





\*\*\*\*\*  
 Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd  
 \*\*\*\*\*

Mparch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
 which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:27:21 2000; MasPar time 166.66 Seconds  
 891.071 Million cell updates/sec

Tabular output not generated.

Title: >US-09-332-866-1  
 Description: (1-25) from US09332866.pep  
 Perfect Score: 316  
 N.A. Sequence: 1 GARGARTTYTNACNCNA.....AYGAYGTNTGYGCNCARGTN 75  
 Comp: CTCTYARRRANTGNGNTT.....TRCTRCANACRCNGNTTCAN

Scoring table: TABLE bktranslated2  
 Gap 40

Nmatch STD : Dbase 0: Query 0

Searched: 1741034 segs, 990042266 bases x 2

Post-processing: Minimum Match 0%  
 Listing first 45 summaries  
 Maximum DB seq length 25

Database: n-pending2  
 1:PCT 2:U6000 3:U6001A 4:U6001B 5:U6002A 6:U6002B  
 7:U6002C 8:U6003A 9:U6003B 10:U6004A 11:U6004B 12:U6004C  
 13:U6005 14:U6006 15:U6007 16:U6008A 17:U6008B 18:U6008C  
 19:U6009A 20:U6009B 21:U6010A 22:U6010B 23:U6011A  
 24:U6011B 25:U6012A 26:U6012B 27:U6013A 28:U6013B  
 29:U6014A 30:U6014B 31:U6014C 32:U6015 33:U6016A  
 34:U6016B 35:U6016C 36:U6017A 37:U6017B 38:U6017C  
 39:U6018A 40:U6018B 41:U6018C 42:U6019 43:NEWP 44:NEWU60

Statistics: Mean 51.092; Variance 101.380; scale 0.504

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB	ID	Description	Pred. No.
-----	-----	-----	-----	-----	-----	-----

No matches found.

Search completed: Fri Jun 23 17:30:12 2000  
 Job time : 171 secs.



WIPERH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Msearch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Fri Jun 23 17:35:51 2000; Maspar time 4610.56 Seconds  
Tabular output not generated. 65.934 Million cell updates/sec

Title: >US-09-332-866-1  
Description: (1-25) from US09332866.pep  
Perfect Score: 316  
N.A. Sequence: 1 GANGARTTYTNACCCNNA.....AYGAYGTNTGYGCNCARGTN 75  
Comp: CTYCTYAARRANTNGNGNTT.....TTCITRCANACRCNGNTYCAN

Scoring table: TABLE bkttranslate2  
Gap 40

Mmatch STD : Dbase 0; Query 0

Searched: 4857316 seqs, 2026511650 bases x 2

Post-processing: Minimum Match 08

Listing first 45 summaries  
Maximum DB seq length 25

Database: emb1-est61  
1:em\_est11 2:em\_est12 3:em\_est13 4:em\_est14 5:em\_est15  
6:em\_est16 7:em\_est17 8:em\_est18 9:em\_est19 10:em\_est20  
11:em\_est21 12:em\_est22 13:em\_est23 14:em\_est24 15:em\_est25  
16:em\_est26 17:em\_est27 18:em\_est28 19:em\_est29 20:em\_est30  
21:em\_est31 22:em\_est32 23:em\_est33 24:em\_est34 25:em\_est35  
26:em\_est36 27:em\_est37 28:em\_est38 29:em\_est39 30:em\_est40  
31:em\_est41 32:em\_est42 33:em\_est43 34:em\_est44 35:em\_est45  
36:em\_est46 37:em\_est47 38:em\_est48 39:em\_est49 40:em\_est50  
41:em\_est51 42:em\_est52 43:em\_est53 44:em\_est54 45:em\_est55  
46:em\_est56 47:em\_est57 48:em\_est58 49:em\_est59 50:em\_est60  
51:em\_est61 52:em\_est62 53:em\_est63 54:em\_est64 55:em\_est65  
56:em\_est66 57:em\_est67 58:em\_est68 59:em\_est69 60:em\_est70  
61:em\_est71 62:em\_est72 63:em\_est73 64:em\_est74 65:em\_est75  
66:em\_est76 67:em\_est77 68:em\_est78 69:em\_est79 70:em\_est80  
71:em\_est81 72:em\_est82 73:em\_est83 74:em\_est84 75:em\_est85  
76:em\_est86 77:em\_est87 78:em\_est88 79:em\_est89 80:em\_est90  
81:em\_est91 82:em\_est92 83:em\_est93 84:em\_est94 85:em\_est95  
86:em\_est96 87:em\_est97 88:em\_est98 89:em\_est99 90:em\_est100  
91:em\_est101 92:em\_est102 93:em\_est103 94:em\_est104 95:em\_est105  
96:em\_est106 97:em\_est107 98:em\_est108 99:em\_est109 100:em\_est110  
101:em\_est111 102:em\_est112 103:em\_est113 104:em\_est114 105:em\_est115  
106:em\_est116 107:em\_est117 108:em\_est118 109:em\_est119 110:em\_est120  
111:em\_est121 112:em\_est122 113:em\_est123 114:em\_est124 115:em\_est125  
116:em\_est126 117:em\_est127 118:em\_est128 119:em\_est129 120:em\_est130  
121:em\_est131 122:em\_est132 123:em\_est133 124:em\_est134 125:em\_est135  
126:em\_est136 127:em\_est137 128:em\_est138 129:em\_est139 130:em\_est140  
131:em\_est141 132:em\_est142 133:em\_est143 134:em\_est144 135:em\_est145  
136:em\_est146 137:em\_est147 138:em\_est148 139:em\_est149 140:em\_est150  
141:em\_est151 142:em\_est152 143:em\_est153 144:em\_est154 145:em\_est155  
146:em\_est156 147:em\_est157 148:em\_est158 149:em\_est159 150:em\_est160  
151:em\_est161 152:em\_est162 153:em\_est163 154:em\_est164 155:em\_est165  
156:em\_est166 157:em\_est167 158:em\_est168 159:em\_est169 160:em\_est170  
161:em\_est171 162:em\_est172 163:em\_est173 164:em\_est174 165:em\_est175  
166:em\_est176 167:em\_est177 168:em\_est178 169:em\_est179 170:em\_est180  
171:em\_est181 172:em\_est182 173:em\_est183 174:em\_est184 175:em\_est185  
176:em\_est186 177:em\_est187 178:em\_est188 179:em\_est189 180:em\_est190  
181:em\_est191 182:em\_est192 183:em\_est193 184:em\_est194 185:em\_est195  
186:em\_est196 187:em\_est197 188:em\_est198 189:em\_est199 190:em\_est200  
191:em\_est201 192:em\_est202 193:em\_est203 194:em\_est204 195:em\_est205  
196:em\_est206 197:em\_est207 198:em\_est208 199:em\_est209 200:em\_est210  
201:em\_est211 202:em\_est212 203:em\_est213 204:em\_est214 205:em\_est215  
206:em\_est216 207:em\_est217 208:em\_est218 209:em\_est219 210:em\_est220  
211:em\_est221 212:em\_est222 213:em\_est223 214:em\_est224 215:em\_est225  
216:em\_est226 217:em\_est227 218:em\_est228 219:em\_est229 220:em\_est230  
221:em\_est231 222:em\_est232 223:em\_est233 224:em\_est234 225:em\_est235  
226:em\_est236 227:em\_est237 228:em\_est238 229:em\_est239 230:em\_est240  
231:em\_est241 232:em\_est242 233:em\_est243 234:em\_est244 235:em\_est245  
236:em\_est246 237:em\_est247 238:em\_est248 239:em\_est249 240:em\_est250  
241:em\_est251 242:em\_est252 243:em\_est253 244:em\_est254 245:em\_est255  
246:em\_est256 247:em\_est257 248:em\_est258 249:em\_est259 250:em\_est260  
251:em\_est261 252:em\_est262 253:em\_est263 254:em\_est264 255:em\_est265  
256:em\_est266 257:em\_est267 258:em\_est268 259:em\_est269 260:em\_est270  
261:em\_est271 262:em\_est272 263:em\_est273 264:em\_est274 265:em\_est275  
266:em\_est276 267:em\_est277 268:em\_est278 269:em\_est279 270:em\_est280  
271:em\_est281 272:em\_est282 273:em\_est283 274:em\_est284 275:em\_est285  
276:em\_est286 277:em\_est287 278:em\_est288 279:em\_est289 280:em\_est290  
281:em\_est291 282:em\_est292 283:em\_est293 284:em\_est294 285:em\_est295  
286:em\_est296 287:em\_est297 288:em\_est298 289:em\_est299 290:em\_est300  
291:em\_est301 292:em\_est302 293:em\_est303 294:em\_est304 295:em\_est305  
296:em\_est306 297:em\_est307 298:em\_est308 299:em\_est309 300:em\_est310  
301:em\_est311 302:em\_est312 303:em\_est313 304:em\_est314 305:em\_est315  
306:em\_est316 307:em\_est317 308:em\_est318 309:em\_est319 310:em\_est320  
311:em\_est321 312:em\_est322 313:em\_est323 314:em\_est324 315:em\_est325  
316:em\_est326 317:em\_est327 318:em\_est328 319:em\_est329 320:em\_est330  
321:em\_est331 322:em\_est332 323:em\_est333 324:em\_est334 325:em\_est335  
326:em\_est336 327:em\_est337 328:em\_est338 329:em\_est339 330:em\_est340  
331:em\_est341 332:em\_est342 333:em\_est343 334:em\_est344 335:em\_est345  
336:em\_est346 337:em\_est347 338:em\_est348 339:em\_est349 340:em\_est350  
341:em\_est351 342:em\_est352 343:em\_est353 344:em\_est354 345:em\_est355  
346:em\_est356 347:em\_est357 348:em\_est358 349:em\_est359 350:em\_est360  
351:em\_est361 352:em\_est362 353:em\_est363 354:em\_est364 355:em\_est365  
356:em\_est366 357:em\_est367 358:em\_est368 359:em\_est369 360:em\_est370  
361:em\_est371 362:em\_est372 363:em\_est373 364:em\_est374 365:em\_est375  
366:em\_est376 367:em\_est377 368:em\_est378 369:em\_est379 370:em\_est380  
371:em\_est381 372:em\_est382 373:em\_est383 374:em\_est384 375:em\_est385  
376:em\_est386 377:em\_est387 378:em\_est388 379:em\_est389 380:em\_est390  
381:em\_est391 382:em\_est392 383:em\_est393 384:em\_est394 385:em\_est395  
386:em\_est396 387:em\_est397 388:em\_est398 389:em\_est399 390:em\_est400  
391:em\_est401 392:em\_est402 393:em\_est403 394:em\_est404 395:em\_est405  
396:em\_est406 397:em\_est407 398:em\_est408 399:em\_est409 400:em\_est410  
401:em\_est411 402:em\_est412 403:em\_est413 404:em\_est414 405:em\_est415  
406:em\_est416 407:em\_est417 408:em\_est418 409:em\_est419 410:em\_est420  
411:em\_est421 412:em\_est422 413:em\_est423 414:em\_est424 415:em\_est425  
416:em\_est426 417:em\_est427 418:em\_est428 419:em\_est429 420:em\_est430  
421:em\_est431 422:em\_est432 423:em\_est433 424:em\_est434 425:em\_est435  
426:em\_est436 427:em\_est437 428:em\_est438 429:em\_est439 430:em\_est440  
431:em\_est441 432:em\_est442 433:em\_est443 434:em\_est444 435:em\_est445  
436:em\_est446 437:em\_est447 438:em\_est448 439:em\_est449 440:em\_est450  
441:em\_est451 442:em\_est452 443:em\_est453 444:em\_est454 445:em\_est455  
446:em\_est456 447:em\_est457 448:em\_est458 449:em\_est459 450:em\_est460  
451:em\_est461 452:em\_est462 453:em\_est463 454:em\_est464 455:em\_est465  
456:em\_est466 457:em\_est467 458:em\_est468 459:em\_est469 460:em\_est470  
461:em\_est471 462:em\_est472 463:em\_est473 464:em\_est474 465:em\_est475  
466:em\_est476 467:em\_est477 468:em\_est478 469:em\_est479 470:em\_est480  
471:em\_est481 472:em\_est482 473:em\_est483 474:em\_est484 475:em\_est485  
476:em\_est486 477:em\_est487 478:em\_est488 479:em\_est489 480:em\_est490  
481:em\_est491 482:em\_est492 483:em\_est493 484:em\_est494 485:em\_est495  
486:em\_est496 487:em\_est497 488:em\_est498 489:em\_est499 490:em\_est500  
491:em\_est501 492:em\_est502 493:em\_est503 494:em\_est504 495:em\_est505  
496:em\_est506 497:em\_est507 498:em\_est508 499:em\_est509 500:em\_est510  
501:em\_est511 502:em\_est512 503:em\_est513 504:em\_est514 505:em\_est515  
506:em\_est516 507:em\_est517 508:em\_est518 509:em\_est519 510:em\_est520  
511:em\_est521 512:em\_est522 513:em\_est523 514:em\_est524 515:em\_est525  
516:em\_est526 517:em\_est527 518:em\_est528 519:em\_est529 520:em\_est530  
521:em\_est531 522:em\_est532 523:em\_est533 524:em\_est534 525:em\_est535  
526:em\_est536 527:em\_est537 528:em\_est538 529:em\_est539 530:em\_est540  
531:em\_est541 532:em\_est542 533:em\_est543 534:em\_est544 535:em\_est545  
536:em\_est546 537:em\_est547 538:em\_est548 539:em\_est549 540:em\_est550  
541:em\_est551 542:em\_est552 543:em\_est553 544:em\_est554 545:em\_est555  
546:em\_est556 547:em\_est557 548:em\_est558 549:em\_est559 550:em\_est560  
551:em\_est561 552:em\_est562 553:em\_est563 554:em\_est564 555:em\_est565  
556:em\_est566 557:em\_est567 558:em\_est568 559:em\_est569 560:em\_est570  
561:em\_est571 562:em\_est572 563:em\_est573 564:em\_est574 565:em\_est575  
566:em\_est576 567:em\_est577 568:em\_est578 569:em\_est579 570:em\_est580  
571:em\_est581 572:em\_est582 573:em\_est583 574:em\_est584 575:em\_est585  
576:em\_est586 577:em\_est587 578:em\_est588 579:em\_est589 580:em\_est590  
581:em\_est591 582:em\_est592 583:em\_est593 584:em\_est594 585:em\_est595  
586:em\_est596 587:em\_est597 588:em\_est598 589:em\_est599 590:em\_est600  
591:em\_est601 592:em\_est602 593:em\_est603 594:em\_est604 595:em\_est605  
596:em\_est606 597:em\_est607 598:em\_est608 599:em\_est609 600:em\_est610  
601:em\_est611 602:em\_est612 603:em\_est613 604:em\_est614 605:em\_est615  
606:em\_est616 607:em\_est617 608:em\_est618 609:em\_est619 610:em\_est620  
611:em\_est621 612:em\_est622 613:em\_est623 614:em\_est624 615:em\_est625  
616:em\_est626 617:em\_est627 618:em\_est628 619:em\_est629 620:em\_est630  
621:em\_est631 622:em\_est632 623:em\_est633 624:em\_est634 625:em\_est635  
626:em\_est636 627:em\_est637 628:em\_est638 629:em\_est639 630:em\_est640  
631:em\_est641 632:em\_est642 633:em\_est643 634:em\_est644 635:em\_est645  
636:em\_est646 637:em\_est647 638:em\_est648 639:em\_est649 640:em\_est650  
641:em\_est651 642:em\_est652 643:em\_est653 644:em\_est654 645:em\_est655  
646:em\_est656 647:em\_est657 648:em\_est658 649:em\_est659 650:em\_est660  
651:em\_est661 652:em\_est662 653:em\_est663 654:em\_est664 655:em\_est665  
656:em\_est666 657:em\_est667 658:em\_est668 659:em\_est669 660:em\_est670  
661:em\_est671 662:em\_est672 663:em\_est673 664:em\_est674 665:em\_est675  
666:em\_est676 667:em\_est677 668:em\_est678 669:em\_est679 670:em\_est680  
671:em\_est681 672:em\_est682 673:em\_est683 674:em\_est684 675:em\_est685  
676:em\_est686 677:em\_est687 678:em\_est688 679:em\_est689 680:em\_est690  
681:em\_est691 682:em\_est692 683:em\_est693 684:em\_est694 685:em\_est695  
686:em\_est696 687:em\_est697 688:em\_est698 689:em\_est699 690:em\_est700  
691:em\_est701 692:em\_est702 693:em\_est703 694:em\_est704 695:em\_est705  
696:em\_est706 697:em\_est707 698:em\_est708 699:em\_est709 700:em\_est710  
701:em\_est711 702:em\_est712 703:em\_est713 704:em\_est714 705:em\_est715  
706:em\_est716 707:em\_est717 708:em\_est718 709:em\_est719 710:em\_est720  
711:em\_est721 712:em\_est722 713:em\_est723 714:em\_est724 715:em\_est725  
716:em\_est726 717:em\_est727 718:em\_est728 719:em\_est729 720:em\_est730  
721:em\_est731 722:em\_est732 723:em\_est733 724:em\_est734 725:em\_est735  
726:em\_est736 727:em\_est737 728:em\_est738 729:em\_est739 730:em\_est740  
731:em\_est741 732:em\_est742 733:em\_est743 734:em\_est744 735:em\_est745  
736:em\_est746 737:em\_est747 738:em\_est748 739:em\_est749 740:em\_est750  
741:em\_est751 742:em\_est752 743:em\_est753 744:em\_est754 745:em\_est755  
746:em\_est756 747:em\_est757 748:em\_est758 749:em\_est759 750:em\_est760  
751:em\_est761 752:em\_est762 753:em\_est763 754:em\_est764 755:em\_est765  
756:em\_est766 757:em\_est767 758:em\_est768 759:em\_est769 760:em\_est770  
761:em\_est771 762:em\_est772 763:em\_est773 764:em\_est774 765:em\_est775  
766:em\_est776 767:em\_est777 768:em\_est778 769:em\_est779 770:em\_est780  
771:em\_est781 772:em\_est782 773:em\_est783 774:em\_est784 775:em\_est785  
776:em\_est786 777:em\_est787 778:em\_est788 779:em\_est789 780:em\_est790  
781:em\_est791 782:em\_est792 783:em\_est793 784:em\_est794 785:em\_est795  
786:em\_est796 787:em\_est797 788:em\_est798 789:em\_est799 790:em\_est800  
791:em\_est801 792:em\_est802 793:em\_est803 794:em\_est804 795:em\_est805  
796:em\_est806 797:em\_est807 798:em\_est808 799:em\_est809 800:em\_est810  
801:em\_est811 802:em\_est812 803:em\_est813 804:em\_est814 805:em\_est815  
806:em\_est816 807:em\_est817 808:em\_est818 809:em\_est819 810:em\_est820  
811:em\_est821 812:em\_est822 813:em\_est823 814:em\_est824 815:em\_est825  
816:em\_est826 817:em\_est827 818:em\_est828 819:em\_est829 820:em\_est830  
821:em\_est831 822:em\_est832 823:em\_est833 824:em\_est834 825:em\_est835  
826:em\_est836 827:em\_est837 828:em\_est838 829:em\_est839 830:em\_est840  
831:em\_est841 832:em\_est842 833:em\_est843 834:em\_est844 835:em\_est845  
836:em\_est846 837:em\_est847 838:em\_est848 839:em\_est849 840:em\_est850  
841:em\_est851 842:em\_est852 843:em\_est853 844:em\_est854 845:em\_est855  
846:em\_est856 847:em\_est857 848:em\_est858 849:em\_est859 850:em\_est860  
851:em\_est861 852:em\_est862 853:em\_est863 854:em\_est864 855:em\_est865  
856:em\_est866 857:em\_est867 858:em\_est868 859:em\_est869 860:em\_est870  
861:em\_est871 862:em\_est872 863:em\_est873 864:em\_est874 865:em\_est875  
866:em\_est876 867:em\_est877 868:em\_est878 869:em\_est879 870:em\_est880  
871:em\_est881 872:em\_est882 873:em\_est883 874:em\_est884 875:em\_est885  
876:em\_est886 877:em\_est887 878:em\_est888 879:em\_est889 880:em\_est890  
881:em\_est891 882:em\_est892 883:em\_est893 884:em\_est894 885:em\_est895  
886:em\_est896 887:em\_est897 888:em\_est898 889:em\_est899 890:em\_est900  
891:em\_est901 892:em\_est902 893:em\_est903 894:em\_est904 895:em\_est905  
896:em\_est906 897:em\_est907 898:em\_est908 899:em\_est909 900:em\_est910  
901:em\_est911 902:em\_est912 903:em\_est913 904:em\_est914 905:em\_est915  
906:em\_est916 907:em\_est917 908:em\_est918 909:em\_est919 910:em\_est920  
911:em\_est921 912:em\_est922 913:em\_est923 914:em\_est924 915:em\_est925  
916:em\_est926 917:em\_est927 918:em\_est928 919:em\_est929 920:em\_est930  
921:em\_est931 922:em\_est932 923:em\_est933 924:em\_est934 925:em\_est935  
926:em\_est936 927:em\_est937 928:em\_est938 929:em\_est939 930:em\_est940  
931:em\_est941 932:em\_est942 933:em\_est943 934:em\_est944 935:em\_est945  
936:em\_est946 937:em\_est947 938:em\_est948 939:em\_est949 940:em\_est950  
941:em\_est951 942:em\_est952 943:em\_est953 944:em\_est954 945:em\_est955  
946:em\_est956 947:em\_est957 948:em\_est958 949:em\_est959 950:em\_est960  
951:em\_est961 952:em\_est962 953:em\_est963 954:em\_est964 955:em\_est965  
956:em\_est966 957:em\_est967 958:em\_est968 959:em\_est969 960:em\_est970  
961:em\_est971 962:em\_est972 963:em\_est973 964:em\_est974 965:em\_est975  
966:em\_est976 967:em\_est977 968:em\_est978 969:em\_est979 970:em\_est980  
971:em\_est981 972:em\_est982 973:em\_est983 974:em\_est984 975:em\_est985  
976:em\_est986 977:em\_est987 978:em\_est988 979:em\_est989 980:em\_est990  
981:em\_est991 982:em\_est992 983:em\_est993 984:em\_est994 985:em\_est995  
986:em\_est996 987:em\_est997 988:em\_est998 989:em\_est999 990:em\_est1000  
991:em\_est1001 992:em\_est1002 993:em\_est1003 994:em\_est1004 995:em\_est1005  
996:em\_est1006 997:em\_est1007 998:em\_est1008 999:em\_est1009 1000:em\_est1010  
1001:em\_est1011 1002:em\_est1012 1003:em\_est1013 1004:em\_est1014 1005:em\_est1015  
1006:em\_est1016 1007:em\_est1017 1008:em\_est1018 1009:em\_est1019 1010:em\_est1020  
1011:em\_est1021 1012:em\_est1022 1013:em\_est1023 1014:em\_est1024 1015:em\_est1025  
1016:em\_est1026 1017:em\_est1027 1018:em\_est1028 1019:em\_est1029 1020:em\_est1030  
1021:em\_est1031 1022:em\_est1032 1023:em\_est1033 1024:em\_est1034 1025:em\_est1035  
1026:em\_est1036 1027:em\_est1037 1028:em\_est1038 1029:em\_est1039 1030:em\_est1040  
1031:em\_est1041 1032:em\_est1042 1033:em\_est1043 1034:em\_est1044 1035:em\_est1045  
1036:em\_est1046 1037:em\_est1047 1038:em\_est1048 1039:em\_est1049 1040:em\_est1050  
1041:em\_est1051 1042:em\_est1052 1043:em\_est1053 1044:em\_est1054 1045:em\_est1055  
1046:em\_est1056 1047:em\_est1057 1048:em\_est1058 1049:em\_est1059 1050:em\_est1060  
1051:em\_est1061 1052:em\_est1062 1053:em\_est1063 1054:em\_est1064 1055:em\_est1065  
1056:em\_est1066 1057:em\_est1067 1058:em\_est1068 1059:em\_est1069 1060:em\_est1070  
1061:em\_est1071 1062:em\_est1072 1063:em\_est1073 1064:em\_est1074 1065:em\_est1075  
1066:em\_est1076 1067:em\_est1077 1068:em\_est1078 1069:em\_est1079 1070:em\_est1080  
1071:em\_est1081 1072:em\_est1082 1073:em\_est1083 1074:em\_est1084 1075:em\_est1085  
1076:em\_est1086 1077:em\_est1087 1078:em\_est1088 1079:em\_est1089 1080:em\_est1090  
1081:em\_est1091 1082:em\_est1092 1083:em\_est1093 1084:em\_est1094 1085:em\_est1095  
1086:em\_est1096 1087:em\_est1097 1088:em\_est1098 1089:em\_est1099 1090:em\_est1100  
1091:em\_est1101 1092:em\_est1102 1093:em\_est1103 1094:em\_est1104 1095:em\_est1105  
1096:em\_est1106 1097:em\_est1107 1098:em\_est1108 1099:em\_est1109 1100:em\_est1110  
1101:em\_est1111 1102:em\_est1112 1103:em\_est1113 1104:em\_est1114 1105:em\_est1115  
1106:em\_est1116 1107:em\_est1117 1108:em\_est1118 1109:em\_est1119 1110:em\_est1120  
1111:em\_est1121 1112:em\_est1122 1113:em\_est1123 1114:em\_est1124 1115:em\_est1125  
1116:em\_est1126 1117:em\_est1127 1118:em\_est1128 1119:em\_est11

